

From: Fredman, Jeffrey  
Sent: Tuesday, January 25, 2005 6:05 AM  
To: STIC-Biotech/ChemLib  
Cc: Dunston, Jennifer  
Subject: FW: Sequence Search 10/826523

PLEASE RUSH.

I Approve.

Jeff Fredman

-----Original Message-----

From: Dunston, Jennifer  
Sent: Monday, January 24, 2005 12:46 PM  
To: Fredman, Jeffrey  
Subject: Sequence Search 10/826523

Jeff,  
Please RUSH this search. The length of SEQ ID NO: 40 is 707 nucleotides. The length of SEQ ID NO: 41 is 3662 nucleotides. The sequences are related in that SEQ ID NO: 41 comprises the sequence of SEQ ID NO: 40.  
Jenn

Please do a sequence search for the nucleic acid sequence of SEQ ID NO: 40 and SEQ ID NO: 41 against the commercial and interference nucleotide databases.  
Thank you.

Jennifer Dunston, Ph.D.  
USPTO Art Unit 1636  
REM 2B76  
Mailbox: REM 2C70  
(571) 272-2916

\*\*\*\*\*

STAFF USE ONLY

Searcher: Jan  
Searcher Phone: 2- 2504  
Date Searcher Picked up: 1/25/05  
Date Completed: 1/27/05  
Searcher Prep/Rev. Time: 10  
Online Time: 10

\*\*\*\*\*

Type of Search

NA Sequence: # ✓  
AA Sequence: #         
Structure: #         
Bibliographic:         
Litigation:         
Patent Family:         
Other:       

\*\*\*\*\*

Vendors and cost where applicable

STN:         
DIALOG:         
QUESTEL/ORBIT:         
LEXIS/NEXIS:         
SEQUENCE SYSTEM: ✓  
WWW/Internet:         
Other(Specify):

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 25, 2005, 23:27:35 ; Search time 15598.5 Seconds  
(without alignments)  
11102.026 Million cell updates/sec

Title: US-10-826-523-41

Perfect score: 3662  
Sequence: 1 ctaaatgtgaagcgttaata.....attccccgaagtgccac 3662

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl: \*  
1: gb\_ba: \*  
2: gb\_hcg: \*  
3: gb\_in: \*  
4: gb\_ov: \*  
5: gb\_ov: \*  
6: gb\_ov: \*  
7: gb\_ph: \*  
8: gb\_pl: \*  
9: gb\_pr: \*  
10: gb\_ro: \*  
11: gb\_sts: \*  
12: gb\_sv: \*  
13: gb\_un: \*  
14: gb\_vt: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	2654.8	72.8	3637	6	AX592968 Sequence
C 2	2656.4	72.5	3637	6	AX592967 Sequence
C 3	2650.6	72.4	3626	12	ASPGREEM1 Y09373 Artificial
4	2644.8	72.2	3624	12	XXU43957 A20702 PKSELS DNA
5	2639	72.1	3681	6	A20702 115651 Sequence 6
6	2639	72.1	3681	6	A20700 115649 Sequence 4
7	2623.6	71.6	5314	6	A20700 pMTVI DNA s
8	2623.6	71.6	5314	6	115649 Sequence 4
C 9	2610.8	71.3	3552	12	AF171227 Shuttle v
C 10	2609.4	71.3	4001	6	AX306330 Sequence
C 11	2558.8	69.9	5540	12	AF179627 Expression
C 12	2550.6	69.3	3633	12	ASPGREEM2 Y09374 Artificial
C 13	2537.8	68.8	3918	6	CQ768840 Sequence
C 14	2517.8	68.8	3918	6	CQ802124 Sequence
C 15	2514.6	67.9	4141	12	XXU5129 C0802124 Sequence
C 16	2485.8	67.9	4141	12	AY190524 Cloning v
C 17	2484.6	67.6	5277	6	AX306326 Sequence
C 18	2474.2	67.4	5277	6	BD244458 Genome an
C 19	2469.4	67.4	5277	6	AR340565 Sequence

20	2451.8	67.0	3928	6	AX591027 Sequence
21	2451.8	67.0	3928	6	AX591180 Sequence
22	2451.8	67.0	3928	6	AX717603 Sequence
23	2449.8	66.9	7933	12	AF359244 Integrati
24	2448.8	66.9	3956	6	A84776 Sequence 21
25	2448.8	66.9	3956	6	AR300179 Sequence
26	2448.8	66.9	3956	6	BD058066 Method fo
27	2446.8	66.8	3715	6	AX306328 Sequence
28	2434	66.5	3970	12	CVE403983 Cloning v
C 29	2420.6	65.9	3274	12	XXU43956 Expression
C 30	2414.6	65.9	3978	12	CVU41513 3' terminal
C 31	2401	65.4	3733	12	AF146395 Cloning v
C 32	2396	65.4	3307	12	SYNPEACX D50399 Unidentific
C 33	2382.6	65.1	4591	12	SYNPEZ18A Cloning vec
C 34	2380.4	65.0	6624	12	CVU37458 Yeast CUP1
C 35	2372.2	64.8	4239	12	AVE401048 Artificial
C 36	2359	64.4	5781	12	CVU37457 Yeast CUP1
C 37	2357.4	64.4	4051	12	XXU5126 Yeast CUP1
C 38	2351	64.2	5275	12	CVU37456 Yeast CUP1
C 39	2349.8	64.2	3230	12	AY663843 Cloning v
C 40	2340.2	63.9	3230	12	AY663844 Cloning v
C 41	2333.4	63.7	5436	12	AY303166 CFP Integ
C 42	2332.8	63.7	5436	12	AY303167 YFP Integ
C 43	2328.6	63.6	3204	12	AY663845 Cloning v
C 44	2327.2	63.5	5385	12	IVE250107 Integrati
C 45	2322.6	63.4	5445	12	IVE306911 Integrati

ALIGNMENTS

RESULT 1  
AX592968/c  
LOCUS AX592968 Sequence 4 from Patent WO02079468.  
DEFINITION AX592968  
ACCESSION AX592968  
VERSION AX592968.1 GI:27950812  
KEYWORDS  
SOURCE  
ORGANISM  
synthetic construct  
synthetic construct  
artificial sequences.  
REFERENCE  
1 Padgett,H.S., Fitzmaurice,W.P. and Lindo,J.A.  
AUTHORS  
TITLE A method of increasing complementarity in a heteroduplex  
JOURNAL Patent: WO 02079468-A 4 10-OCT-2002;  
Large Scale Biology Corporation (US)  
FEATURES  
source  
1. 3637  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"

ORIGIN

Query Match	72.8%	Score 2664.8;	DB 6;	Length 3637;
Best Local Similarity	83.8%	Pred. No. 0;		
Matches 3070;	Conservative 0;	Mismatches 567;	Indels 25;	Gaps 4;
Cy 1	CTAAATGTGAAGCGTTAATATTTTGTAAATTCGCGTTAAATTTTGTAAATCAGCTC	60		
Db 3637	CTAAATGTGAAGCGTTAATATTTTGTAAATTCGCGTTAAATTTTGTAAATCAGCTC	3578		
Cy 61	ATTTTAAACCAATAGCGCCGAATCGCAAAATCCCTTAATATCAAAAGATAGACCGA	120		
Db 3577	ATTTTAAACCAATAGCGCCGAATCGCAAAATCCCTTAATATCAAAAGATAGACCGA	3518		
Cy 121	GATGGGTGAGTGTTCCTTCCAGTTTGGAACAAGTCCACTTAAGAAGCGGAGCTC	180		
Db 3517	GATGGGTGAGTGTTCCTTCCAGTTTGGAACAAGTCCACTTAAGAAGCGGAGCTC	3458		
Cy 181	CAACGTAAAGGCGGAAACCGCTCTATCAGGCGATGCGCCACTAGTGAACCATGACC	240		
Db 3457	CAACGTAAAGGCGGAAACCGCTCTATCAGGCGATGCGCCACTAGTGAACCATGACC	3398		

Qy	241	CTATCAAGTTTTTGGGGGTGAGGTGCTTAAGAACAATAATCGAACCTTAAGGAG	300
Db	3397	CTATCAAGTTTTTGGGGGTGAGGTGCTTAAGAACAATAATCGAACCTTAAGGAG	333
Qy	301	CCCCGATTTAGAGCTTGACGGGAAAAGCCGGGAAACGTGGCGAAGAAAGAAAGGAA	360
Db	3337	CCCCGATTTAGAGCTTGACGGGAAAAGCCGGGAAACGTGGCGAAGAAAGAAAGGAA	327
Qy	361	AGCGAAAGAGCGGGCGCTTAGGGCGCTGGCAAGTATAGCGATCACGTGCGGTAAAC	420
Db	3277	AGCGAAAGAGCGGGCGCTTAGGGCGCTGGCAAGTATAGCGATCACGTGCGGTAAAC	321
Qy	421	CACACCGCGCGCTTAATGCGCGCGCTAACAGGCGCGGTGCCATTAGAGCTGG	480
Db	3217	CACACCGCGCGCTTAATGCGCGCGCTAACAGGCGCGGTGCCATTAGAGCTGG	315
Qy	481	CAACTGTGGGAAAGGCGGATCGGTGCGGGCTTTCGTATTACCCAGCTGGCGAAG	540
Db	3157	CAACTGTGGGAAAGGCGGATCGGTGCGGGCTTTCGTATTACCCAGCTGGCGAAG	309
Qy	541	GGATGTGCTGCAAGGCGATTAAAGTGGGTAAACCCAGGGTTTTCCAGTCAGAGTGG	600
Db	3097	GGATGTGCTGCAAGGCGATTAAAGTGGGTAAACCCAGGGTTTTCCAGTCAGAGTGG	303
Qy	601	TAAACGACGGCCAGTGAAGCGCGCTCTTCAATTCAGTTTTTGAACCCGAGAGACGG	660
Db	3037	TAAACGACGGCCAGTGAAGCGCGCTCTTCAATTCAGTTTTTGAACCCGAGAGACGG	297
Qy	661	GCAGACTCGCGGGCAAAATGTGTTTTACAGCTGATGAGACAGATGAAGATCTCGAC	720
Db	2977	CGCGGTGGCG-----CGCTTGAAGCTAGTGATCCCC--CGG	293
Qy	721	GCTGCGAAACACGCACTAGATTAACCTAGAAAGATATATATGTGACGTACGTAA	780
Db	2938	GCTGCGAAATCTTATTTATTTATGTTATCATCATCCATGTGTATCCAGACAGTTA	287
Qy	781	AGATATCATGGGTAAATTAACGATGGGATCTGTAAATACACTCATATAGGCGCAAT	840
Db	2878	CAAACTCAAGAGACCAATGTGTACACCTTTCGTTGGGATCTTTGAAAGGCGCAATT	281
Qy	841	TGGGTACCGGGCCCCCTCGAGGTGACGCGGATCGAATAACCTGATATCGAATCTTCG	900
Db	2818	GTGTGACAGGTAAATGTGTGTCTGTAAAGAACAGGCCATCGCCAAATGAGATATTT	275
Qy	901	AACCCGGGGATCCACTGTGTTCTAGAGCGCGCCACCGCGGTGAGCTCAGCTTTGT	960
Db	2788	GTTGATTAATGTCTGTGTGAACGGAATCCATCTTAATGTGTGGCAATTTGAAGT	269
Qy	961	TCCCTTAGTGGGTAAATAGATCCCATGCGCAATTTTAACCACTATCTTTAG	1020
Db	2698	TAGCTTTAGTCCATCTTTTGTGTGTCTGCGGAGTGAATCAATGTGTAG--TTATA	2640
Qy	1021	GATTATCTAGCTGCATCAGATCAATCGTGGGTCTTTTTCCGGCTCAGTCATCGCC	1080
Db	2639	TTGTACTGAGATTTGTGTGTCGAGATGTTTTCATCTCTTTAAATAATCAATACCTTTAAC	2580
Qy	1081	CAAGCTGCGCTATCTGGGCAATCGGGGAGAGAAAGCCGTGCTTTTCCCGAGGTTG	1140
Db	2579	TCGATACGATTAAAGAGGTATCACTTTCAAACTTGCATTCACACGCGCTTGTGTTC	2520
Qy	1141	AAGCGGACGTGAAGAGTTGCGAGAGATGACTGTGCTCATTTGACGTTGACGAAAC	1200
Db	2519	CCGTATCTTTGAAGAATATAGTCGTTCCGTATCATTAACCTTCGGGACATGGACCTTG	2460
Qy	1201	GCAGGTTTCCATGATGATTCGGGAAGGTGTGCGCATGACGCTTTAAAGGTGAAC	1260
Db	2459	AAAAAGTATGCGCTTTCATATGATCGGATTAACGGAAAGATTTGAACACCATTAAG	2400
Qy	1261	TCGTTACGCGCACTTGGGATACAGTTGTCGCGGCTTTTCCGACACAGTTCCGATGG	1320
Db	2399	AAAGTATGACAAAGTGTGGCAATGAAACAGATATGTTTTCAAGTATGTGCAATTAATTA	2340
Qy	1321	TCAGCCGGAAGCGATACGAACCCGAACAATAACCGGACAGCCGGAACTGCGTCCG	1380

Db	2339	AGGGTAAGCTTTCCGTAATGTAAGCATCACTTCAACCTCTCCACTGACAGAAAAATTGTGC	228
Qy	1381	GTGTGCAAGATTATATGACAGCGGTGCGCGCTGGAGATTATACGTACAGCAGACGGGATAC	144
Db	2279	CCATTACATCAACCATTAATTATTCACAAAGAAATTGGGACAACTCAGTAGAAAAATTCTTCT	222
Qy	1441	CTGGCTGAGATGCGCGAANAATGACATGATACCCCTGAGTTAACCCGCGCGCGCTT	150
Db	2219	C--CTTTACTCATCGGTATCCAGACTTTTGTCCCTTTAGTAGGGTTAATTCGCGGCTT	216
Qy	1501	GGCGTAATCATNGCTCAGCTGTGTTCTGTGTAAATGTGTATCCGCTCAATTCACAA	156
Db	2162	GGCGTAATCATNGCTCAGCTGTGTTCTGTGTAAATGTGTATCCGCTCAATTCACAA	210
Qy	1561	CAACATPCAGGCCGGGAAGCATTAAGTGTAAAGCTCGGGTGCCTAATGAGTAGCTTAAT	162
Db	2102	CAACATPCAGGCCGGGAAGCATTAAGTGTAAAGCTCGGGTGCCTAATGAGTAGCTTAAT	204
Qy	1621	CACATTAAATGGCGTGGGCTCACTGCGCGTTTCCAGTGGGGAACCTGTGCGTCCAGCT	168
Db	2042	CACATTAAATGGCGTGGGCTCACTGCGCGTTTCCAGTGGGGAACCTGTGCGTCCAGCT	198
Qy	1681	GCATTAAATGATGCGCCAAACGCGCGGGAGAGCGCGGTTTGCGTATTGGCGCTCTTCCGC	174
Db	1982	GCATTAAATGATGCGCCAAACGCGCGGGAGAGCGCGGTTTGCGTATTGGCGCTCTTCCGC	192
Qy	1741	TTCTCGCTCACTGACCTGCGTGGGCTCGGTGCTGGCGTGGCGGACGGGTATCAGCTCA	180
Db	1922	TTCTCGCTCACTGACCTGCGTGGGCTCGGTGCTGGCGTGGCGGACGGGTATCAGCTCA	186
Qy	1801	CTCAAGGCGGTATACGGTTATCCACAGAAATCAGGGGATTAACGACGAAAGAAACATGTG	186
Db	1862	CTCAAGGCGGTATACGGTTATCCACAGAAATCAGGGGATTAACGACGAAAGAAACATGTG	180
Qy	1861	AGCAAAAGGCGAGCAAAAGGCCAGGAACCGTTAAAGCCGCGTGTGCTGGCGTTTTCACA	1920
Db	1802	AGCAAAAGGCGAGCAAAAGGCCAGGAACCGTTAAAGCCGCGTGTGCTGGCGTTTTCACA	174
Qy	1921	TAGGCTCCGCCCCCTGACGAGCATCACAAAATATGACGCTCAAGTCAAGGTGGCGAAA	198
Db	1742	TAGGCTCCGCCCCCTGACGAGCATCACAAAATATGACGCTCAAGTCAAGGTGGCGAAA	168
Qy	1981	CCCGACAGGACTATTAAGAATACAGGCGTTTCCCTGAAAGTCCCTGATGCGCTCTCC	2040
Db	1682	CCCGACAGGACTATTAAGAATACAGGCGTTTCCCTGAAAGTCCCTGATGCGCTCTCC	1622
Qy	2041	TGTTCCGACCCCTGACGAGTACCTGTGCGGCTTTGCCCTTGGGAAAGGTGAC	2100
Db	1622	TGTTCCGACCCCTGACGAGTACCTGTGCGGCTTTGCCCTTGGGAAAGGTGAC	1563
Qy	2101	GCTTTCTCATAGCTCAGCTGAGGTATCTCAGTTCCGTGTAGTGTGTTCCGCTCAAGCT	2160
Db	1562	GCTTTCTCATAGCTCAGCTGAGGTATCTCAGTTCCGTGTAGTGTGTTCCGCTCAAGCT	1503
Qy	2161	GGGCTGTGTGACGAAACCCCGTTACGCCGCCGACCGCTGCGGCTTTATCCGTTAATCTATCG	2220
Db	1502	GGGCTGTGTGACGAAACCCCGTTACGCCGCCGACCGCTGCGGCTTTATCCGTTAATCTATCG	1443
Qy	2221	TCTTGAAGTCCAAACCGGTAGACACGATTATGCGCCGACCGCTGCGGCTTTATCCGTTAATCTATCG	2280
Db	1442	TCTTGAAGTCCAAACCGGTAGACACGATTATGCGCCGACCGCTGCGGCTTTATCCGTTAATCTATCG	1383
Qy	2281	GATTAGACAGACCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGTGGCTTAACCTA	2340
Db	1382	GATTAGACAGACCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGTGGCTTAACCTA	1323
Qy	2341	CGGCTACACTAGAGACAGATTATTTGGTATCTGGGCTCTGTCTAAAGCAATTCCTTCGG	2400
Db	1322	CGGCTACACTAGAGACAGATTATTTGGTATCTGGGCTCTGTCTAAAGCAATTCCTTCGG	1263
Qy	2401	AAAAAGATTGTAGCTCTTGATCCGGCAACAACACCGCTGTATGCGGTGTTTTTTT	2460



Db 1262 AAAAGAGTTGGTAGCTCTTGATCCGGCAACAAACACCGCTGGTAGCGGTGTTT 1203  
Qy 2461 TGTTCGACAGACAGATTTAGCGCAGAAAAAGATCTCAAGAAAGATCTTTGATCTT 2520  
Db 1202 TGTTCGACAGACAGATTTAGCGCAGAAAAAGATCTCAAGAAAGATCTTTGATCTT 1143  
Qy 2521 TTTCAGGGGCTGACGCTCAGTGAAGCAAAACACGTTAAGGATTTTGGTCATGAG 2560  
Db 1142 TTTCAGGGGCTGACGCTCAGTGAAGCAAAACACGTTAAGGATTTTGGTCATGAG 1083  
Qy 2581 ATTATCAAAAAAGATCTTCACTAGATCTTTTAAATTAATAATGAATTTAAATCAAT 2640  
Db 1082 ATTATCAAAAAAGATCTTCACTAGATCTTTTAAATTAATAATGAATTTAAATCAAT 1023  
Qy 2641 CTAAAGTATATATAGTAAACTTGGTGTGACAGTTACCAATGCTTAACTAGTGAAGCAC 2700  
Db 1022 CTAAAGTATATATAGTAAACTTGGTGTGACAGTTACCAATGCTTAACTAGTGAAGCAC 963  
Qy 2701 TATCTGACGATCTGTCTATTTGCTTATCCATAGTTGCTGACCTCCCGCTGCTGAT 2760  
Db 962 TATCTGACGATCTGTCTATTTGCTTATCCATAGTTGCTGACCTCCCGCTGCTGAT 903  
Qy 2761 AACTACGATACGGAGGGCTTACCATCTGGCCCAAGCTGCAATGATACCGAGAGACC 2820  
Db 902 AACTACGATACGGAGGGCTTACCATCTGGCCCAAGCTGCAATGATACCGAGAGACC 843  
Qy 2821 ACCGCTACCGGCTCAGATTTTATGCAATTAACAGCCAGCCGAAAGGGCCGAGCCGAG 2880  
Db 842 ACCGCTACCGGCTCAGATTTTATGCAATTAACAGCCAGCCGAAAGGGCCGAGCCGAG 783  
Qy 2881 AAGGATCTGCACTTTATCGGCTCATCAGTCTATTTAATTTGGTCGGGAGGCTAG 2940  
Db 782 AAGGATCTGCACTTTATCGGCTCATCAGTCTATTTAATTTGGTCGGGAGGCTAG 723  
Qy 2941 AGTAAAGTATGTCGACGTTAATAGTTTGGCGCAAGTTGTCCTATGCTACAGGCAATCGT 3000  
Db 722 AGTAAAGTATGTCGACGTTAATAGTTTGGCGCAAGTTGTCCTATGCTACAGGCAATCGT 663  
Qy 3001 GGTGTACGCTCGTGTGTTGGTATGAGGCTTCACTGAGCTCCGTTCCCAACGATCAAGCG 3060  
Db 662 GGTGTACGCTCGTGTGTTGGTATGAGGCTTCACTGAGCTCCGTTCCCAACGATCAAGCG 603  
Qy 3061 AGTAAAGTATGTCGACGTTAATAGTTTGGCGCAAGTTGTCCTATGCTACAGGCAATCGT 3120  
Db 602 AGTAAAGTATGTCGACGTTAATAGTTTGGCGCAAGTTGTCCTATGCTACAGGCAATCGT 543  
Qy 3121 TGTTCAGAAAGTATGTCGACGTTAATAGTTTGGCGCAAGTTGTCCTATGCTACAGGCAATCGT 3180  
Db 542 TGTTCAGAAAGTATGTCGACGTTAATAGTTTGGCGCAAGTTGTCCTATGCTACAGGCAATCGT 483  
Qy 3181 TCTTACTGTCAATGTCATCCGTAAGATGCTTTCTGTGAATGCTGAGTATCTCAACCAAGTC 3240  
Db 482 TCTTACTGTCAATGTCATCCGTAAGATGCTTTCTGTGAATGCTGAGTATCTCAACCAAGTC 423  
Qy 3241 ATTCTGAGAAATAGTATGTCGACGTTAATAGTTTGGCGCAAGTTGTCCTATGCTACAGGCAATCGT 3300  
Db 422 ATTCTGAGAAATAGTATGTCGACGTTAATAGTTTGGCGCAAGTTGTCCTATGCTACAGGCAATCGT 363  
Qy 3301 TACCGCCGCAATAGCAGAACTTTAAAGTGTCTATCATCTTGGGAGCAAAAGGTTCTTGGGGCG 3360  
Db 362 TACCGCCGCAATAGCAGAACTTTAAAGTGTCTATCATCTTGGGAGCAAAAGGTTCTTGGGGCG 303  
Qy 3361 AAAAAGTCTCAAGATCTTTAAGGCTGTTGAGATCAAGTTGATGATCCCACTGTCAC 3420  
Db 302 AAAAAGTCTCAAGATCTTTAAGGCTGTTGAGATCAAGTTGATGATCCCACTGTCAC 243  
Qy 3421 CAACGTGATCTCAAGATCTTTAAGGCTGTTGAGATCAAGTTGATGATCCCACTGTCAC 3480  
Db 242 CAACGTGATCTCAAGATCTTTAAGGCTGTTGAGATCAAGTTGATGATCCCACTGTCAC 183  
Qy 3481 GCAAAATGCGCAAAAGGAAATAGGCGCAACGCAAAAGTGTGAATACTCATACTCTT 3540  
Db 182 GCAAAATGCGCAAAAGGAAATAGGCGCAACGCAAAAGTGTGAATACTCATACTCTT 123

Qy 3541 CCTTTTCAATATATGAAGCAATTTATCAGGGTATTTGTCATGAGCGGATCATATT 3600  
Db 122 CCTTTTCAATATATATGAAGCAATTTATCAGGGTATTTGTCATGAGCGGATCATATT 63  
Qy 3601 TGAATATATTTAGAAAAATTAACAAATAGAGGGTTCGCGCACATTTCCCGAAAGTCC 3660  
Db 62 TGAATATATTTAGAAAAATTAACAAATAGAGGGTTCGCGCACATTTCCCGAAAGTCC 3  
Qy 3661 AC 3662  
Db 2 AC 1  
RESULT 2  
AX592967/c 3637 bp DNA linear PAT 27-JAN-2003  
LOCUS Sequence 3 from Patent WO02079468.  
DEFINITION AX592967  
ACCESSION AX592967  
VERSION AX592967.1 GI:27950811  
KEYWORDS  
SOURCE  
ORGANISM  
synthetic construct  
synthetic construct  
artificial sequences.  
REFERENCE  
1 Padgett,H.S., Fitzmaurice,W.P. and Lindo,J.A.  
AUTHORS A method of increasing complementarity in a heteroduplex  
TITLE Patent: WO 02079468-A 3 10-OCT-2002;  
JOURNAL Large Scale Biology Corporation (US)  
FEATURES  
LOCATION/Qualifiers  
SOURCE 1..3637  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
ORIGIN  
Query Match 72.5%; Score 2656.4; DB 6; Length 3637;  
Best Local Similarity 84.0%; Pred. No. 0;  
Matches 3079; Conservative 0; Mismatches 556; Indels 29; Gaps 6;  
Qy 1 CTAAATGTAAGGCTTAATTTTGTAAATTCGCGTAAATTTTGTAAATCAGCTC 60  
Db 3637 CTAAATGTAAGGCTTAATTTTGTAAATTCGCGTAAATTTTGTAAATCAGCTC 3578  
Qy 61 ATTTTAAACCAATAGGCGCAAAATCGCAAAATCCCTTATTAATCAAAAGATAGACGA 120  
Db 3577 ATTTTAAACCAATAGGCGCAAAATCGCAAAATCCCTTATTAATCAAAAGATAGACGA 3518  
Qy 121 GATAGGTTAGAGTGTGTTTCCAGTTTGGAAACAGAGTCCACTATTAAGAAAGTGGACTC 180  
Db 3517 GATAGGTTAGAGTGTGTTTCCAGTTTGGAAACAGAGTCCACTATTAAGAAAGTGGACTC 3458  
Qy 181 CAACGTCAAGGCGGAAACCGCTCATCAGGGCGATGCGCCACTAGCTAACCATCACC 240  
Db 3457 CAACGTCAAGGCGGAAACCGCTCATCAGGGCGATGCGCCACTAGCTAACCATCACC 3338  
Qy 3577 CTAAATCAAGTTTGTGGGGTTCGAGGTGCTGTAAGCACTAAATCGGAACCTTAAGGGAG 3338  
Db 241 CTAAATCAAGTTTGTGGGGTTCGAGGTGCTGTAAGCACTAAATCGGAACCTTAAGGGAG 300  
Qy 3397 CTAAATCAAGTTTGTGGGGTTCGAGGTGCTGTAAGCACTAAATCGGAACCTTAAGGGAG 3338  
Db 301 CCCCCCATTTTAAAGCTTGAACGGGGAAAGCCGGCGAAAGTGGCGAAGAAAGGAAAGAA 360  
Qy 3337 CCCCCCATTTTAAAGCTTGAACGGGGAAAGCCGGCGAAAGTGGCGAAGAAAGGAAAGAA 3278  
Db 361 AGCGAAAGAGAGGCGCTAGGCGCTGCGCAAGTGTAGCGGTCAAGCTGCGCGTAAACAC 420  
Qy 3277 AGCGAAAGAGAGGCGCTAGGCGCTGCGCAAGTGTAGCGGTCAAGCTGCGCGTAAACAC 3218  
Db 421 CAACCCGCGCGCTTAATGCGCGCTTACAGGGCGCGCTCCATTCGCTACGAGCTGCG 480  
Qy 3217 CAACCCGCGCGCTTAATGCGCGCTTACAGGGCGCGCTCCATTCGCTACGAGCTGCG 3158  
Db 481 CAACGTGTTGGAAAGGCGGATCGGTGCGGGCTCTTCCGCTATTACGCCAGCTGCGGAAG 540

Db 3157 CAACCTGGGAAGGCGATCGGTGCGGCGCTCTTCTGCTAATTACGCGACCTGGCCGAAGG 3098  
QY 541 GGGATGTGCTGCAAGGCGATTAAATTGGGTAAAGCGAGGGTTTCCAGTCAAGAGTTG 600  
Db 3097 GGGATGTGCTGCAAGGCGATTAAATTGGGTAAAGCGAGGGTTTCCAGTCAAGAGTTG 3038  
QY 601 TAAAGAGCGGCGAGGAGCGCGCCCTGTTCAATTACGTTTGAACCCGTGAGAGAGCG 660  
Db 3037 TAAAGAGCGGCGAGGAGCGCGCGCTGAATACACTCATATAGGCGAATTGAGCTCCA 2978  
QY 661 GCAGACTGCGGGTGCMAATGTGTTTACAGCGTGAATGAGCAGATGAAGATGCTGAGAC 720  
Db 2977 CCGCGGTGCGGCGCGCTGAGAACTAGTGAATCCCCCGGCTGCGAGAAATCTTAATTTGT 2918  
QY 721 GCTGCGAAGCAGCGAGCTAGATTAAACCTGAAGAAATATATATATTTGAGCGTAA 780  
Db 2917 ATAGTTCACTCCATGCGCATGTGTATCCAGCGAGCTGTAACTCAAGAAAGACATGT 2858  
QY 781 AGATATATC--ATGCGTAAATTTGACGCAATGGGATCTGTATATACAGCTCATATAGGCGGA 838  
Db 2857 GGTCTCTCTTTCTGTTGGGATCTTTCGAAAGGCGAGATTGTGTGACAGGTATGTGTGT 2798  
QY 839 ATTGGGTACCGGCGCCCCCTTCGAGGTGACGCTATCGATAGCTTGATATCGAATTCCT 898  
Db 2797 CTGTAAAGAGCAGGCGCATCGCAATTGGAGTATTTGTGTATATGTCTGTAGTT 2738  
QY 899 GCAGCCCGGGGGATCCACTAGTTCAGAGCGCGCCGACCGCGGTGAGCTCCGCTTT 958  
Db 2737 GAGCGCTTCACATTTCATATGTGTGTCTAATTTGAAGTTAACTTATTCATCTTTT 2678  
QY 959 GTTCCCTTTAGTGAAGGTTAAATTAGATCCAGTGCATTTTACGAGACTATCTTCT 1018  
Db 2677 GTTGTCTGCGCATGATGATATCATTTGTGTGA- GTTATATAGTTATTTCAATTGTGTCA 2619  
QY 1019 AGGTTATCTAGCTGATCAAGATCATATGTGGGCTTTTTCGGGCTCACTCATCG 1078  
Db 2618 AGAATGTTTCATCTCTTTAAATCAATACCTTTTAATCTGATTC-----TATTAACA 2564  
QY 1079 CCCAAGCTGCGCTATCTGGGATCGGAGAGAAAGAACCCGCTTTTCCGCGAGGT 1138  
Db 2563 GGGATATCACCTTCAAACTTGACTTC-----AGCAGGTCTTGTATGTCCGCT 2515  
QY 1139 TGAAGCGGATGCAAAAGTTTGCAGAGATGACTGTCTGCTGATTAAGCTTGAAGCAAA 1198  
Db 2514 ATCTTTAAATAATATATCTTCTCTGTATCATTAACCTTGGGATGCACTCTTGAATA 2455  
QY 1199 ACGCAGCTTACCATATGATGATTTGGGAAAGTGTGGCATGACAGCCTTTAAAGGTAACT 1258  
Db 2454 GTATGCGCTTCTATGATCTGGGTATCTTGAAGACATTTGAACATTAAGAAAGT 2395  
QY 1259 GTTCTGTCAGGCCACTGGGATACAGTTGTGCGGCTTTTCGGGACAAGTTCCGGAT 1318  
Db 2394 AGTGAACAAGTGTGGCATGGAACAGTATGTTTCCAGTATGCAATTAATTTAAAGGT 2335  
QY 1319 GGTCAAGCCGAACGATCAAGCAACCCGAACATACCGGAGCAGCGGAACTGCGGTGC 1378  
Db 2334 AAGTTTTC-----GATATGTTCATCACTTCACTCTTCCACTGACAGAAATTTGTGC 2280  
QY 1379 CGGTGTGAGATTAATGACAGCGGTGCGGCTGGGATATTAAGTCAAGGAGAGAGCGGTA 1438  
Db 2279 CCATTAACATCACCATCTAATTAACAAGAAATTTGGAGCACTCCAGTGAAGATTTCTCT 2220  
QY 1439 TCCGTGCTGATGCCCGAAGATGACATGATACCCCGTGAATTAACCGGCGGCGCGC 1498  
Db 2219 CCTTAATCATGTGTAACCGAGCTTTGTTCCTTTAGTAGGGTTA-----ATTGCGCGC 2165  
QY 1499 TTGGCTAATCATGCTATAGCTTTCTGTGTGAATTTGTATCCGCTCAATTTCA 1558  
Db 2164 TTGGCTAATCATGCTATAGCTTTCTGTGTGAATTTGTATCCGCTCAATTTCA 2105  
QY 1559 CACCAATATGAGCGCGAAGCAATAAGTAAAGCTGGGCTGCTAATGATGAGCTAA 1618  
|||||

Db 2104 CACAACATAGAGCCGGAAGCATTAAGTTAAAGCCTGGGCTGCTAATGATGAGCTTA 2045  
QY 1619 CTCACTTAATTTGGCTTGCCTCATCTGCCCTTTCAGTCCGGAACCTGTGCTCAG 1678  
Db 2044 CTCACTTAATTTGGCTTGCCTCATCTGCCCTTTCAGTCCGGAACCTGTGCTCAG 1985  
QY 1679 CTGCAATTAATGAAATCCGCGCAAGCGGCGGAGAGGGGCTTTGTGATTTGGCGCTTTCC 1738  
Db 1984 CTGCAATTAATGAAATCCGCGCAAGCGGCGGAGAGGGGCTTTGTGATTTGGCGCTTTCC 1925  
QY 1739 GCTTCTCGCTCACTGACTGCTGCTCGCTCGTGTGCTGCTGCGCGGAGCTGATAGCT 1798  
Db 1924 GCTTCTCGCTCACTGACTGCTGCTGCTCGCTCGTGTGCTGCTGCGGAGCTGATAGCT 1865  
QY 1799 CACTCAAGCGGCTAATAGGTTATCCACAGAAATCGAGGATTAACGAGAAAGAAATG 1858  
Db 1864 CACTCAAGCGGCTAATAGGTTATCCACAGAAATCGAGGATTAACGAGAAAGAAATG 1805  
QY 1859 TGAACAAAAGCCAGCAAAAGCCAGGAACGCTTAAAAAAGCGCGTGTGCGCTTTTC 1918  
Db 1804 TGAACAAAAGCCAGCAAAAGCCAGGAACGCTTAAAAAAGCGCGTGTGCGCTTTTC 1745  
QY 1919 CATAGCTCCGCCCCCTTGAAGCATCAAAATTCAGCTCAATGATGAGTGGCGA 1978  
Db 1744 CATAGCTCCGCCCCCTTGAAGCATCAAAATTCAGCTCAATGATGAGTGGCGA 1685  
QY 1979 AACCCGACAGACTTAAGATACAGAGGCTTTCCCGTGAAGCTCCCTGTGCTCT 2038  
Db 1684 AACCCGACAGACTTAAGATACAGAGGCTTTCCCGTGAAGCTCCCTGTGCTCT 1625  
QY 2039 CCTGTTCGACCTCCCGCTTAACCGGATACCTGTCCGCTTCTTCCGGAAGCTG 2098  
Db 1624 CCTGTTCGACCTCCCGCTTAACCGGATACCTGTCCGCTTCTTCCGGAAGCTG 1565  
QY 2099 GCGCTTTCATAGTCAAGCTGATGATATCTCACTTGTGATGCTGTTCCGCAAG 2158  
Db 1564 GCGCTTTCATAGTCAAGCTGATGATATCTCACTTGTGATGCTGTTCCGCAAG 1505  
QY 2159 CTGGGCTGTGTGACGAACCCCGCTTCAAGCCGACGCTGCGCTTATCCGGAATCAT 2218  
Db 1504 CTGGGCTGTGTGACGAACCCCGCTTCAAGCCGACGCTGCGCTTATCCGGAATCAT 1445  
QY 2219 GCTTGTGATCAACCCGCTTAAGACATTAATGCTTCACTGAGAGCCACTGTATAC 2278  
Db 1444 GCTTGTGATCAACCCGCTTAAGACATTAATGCTTCACTGAGAGCCACTGTATAC 1385  
QY 2279 AGAATTAAGCAGAGGATATGTAGCGGCTCTACAGATTTTGAAGTGTGCTTAAC 2338  
Db 1384 AGAATTAAGCAGAGGATATGTAGCGGCTCTACAGATTTTGAAGTGTGCTTAAC 1325  
QY 2339 TACGCTTACACTGAAGAGACATTTGTATCTGCGCTGTGCTGAAGCAGTTAACTTC 2398  
Db 1324 TACGCTTACACTGAAGAGACATTTGTATCTGCGCTGTGCTGAAGCAGTTAACTTC 1265  
QY 2399 GAAAAGAGTGTGTACTCTTGAATCCGGAACAAACCCGCTGTAGCGGTGTTT 2458  
Db 1264 GAAAAGAGTGTGTACTCTTGAATCCGGAACAAACCCGCTGTAGCGGTGTTT 1205  
QY 2459 TTTGTTTGAAGCAGATTAAGCGGAGAAAAAGAACTCTGAAGAGATCTTTGATC 2518  
Db 1204 TTTGTTTGAAGCAGATTAAGCGGAGAAAAAGAACTCTGAAGAGATCTTTGATC 1145  
QY 2519 TTTTCTTACGGGCTTGAAGCTCATGTTGAAGCAAAACTCACTTAAGGATTTTGTGATG 2578  
Db 1144 TTTTCTTACGGGCTTGAAGCTCATGTTGAAGCAAAACTCACTTAAGGATTTTGTGATG 1085  
QY 2579 AGATTAATCAAAAAGATCTTCACTAGATCTTTTAATTTAAATGAAGTTTAATCA 2638  
Db 1084 AGATTAATCAAAAAGATCTTCACTAGATCTTTTAATTTAAATGAAGTTTAATCA 1025  
QY 2639 ATCTAAATATATATAGATAAATTTGCTGACAGTTAACAAATGCTTAATCAAGTACGCA 2698  
Db 1024 ATCTAAATATATATAGATAAATTTGCTGACAGTTAACAAATGCTTAATCAAGTACGCA 965  
|||||

```

Qy 2699 CCTATCTCAGGATCTGTCTATTTCTTCAATCAGTGGCTGACTCCCGTGTGTAG 2758
Db 964 CTAATCTCAGGATCTGTCTATTTCTTCAATCAGTGGCTGACTCCCGTGTGTAG 905
Qy 2759 ATAATCAGATACGGAGGGCTTACCATCTGGCCCAAGTGTGCATATGACCGGAGAC 2818
Db 904 ATAATCAGATACGGAGGGCTTACCATCTGGCCCAAGTGTGCATATGACCGGAGAC 845
Qy 2819 CCACGCTCACCAGCTCCAGATTTATCAGCAATTAACAGCCAGCGGAGGGCCGAGGC 2878
Db 844 CCACGCTCACCAGCTCCAGATTTATCAGCAATTAACAGCCAGCGGAGGGCCGAGGC 785
Qy 2879 AGAAGTGTCTGTGAATTTATCCGCTCCATCCAGTCTATTAATTTGTGCGGGAAAGCT 2938
Db 784 AGAAGTGTCTGTGAATTTATCCGCTCCATCCAGTCTATTAATTTGTGCGGGAAAGCT 725
Qy 2939 AGAATAGTATGTTGCCAGTTAATAGTTTGGCAACGTTGTGCAATGCTACAGGCATC 2998
Db 724 AGAATAGTATGTTGCCAGTTAATAGTTTGGCAACGTTGTGCAATGCTACAGGCATC 665
Qy 2999 GTGTGTCAAGCTGTGTGTGTGTATGCTTCAATTCAGCTCCGTTCCCAAGATCAAG 3058
Db 664 GTGTGTCAAGCTGTGTGTGTGTATGCTTCAATTCAGCTCCGTTCCCAAGATCAAG 605
Qy 3059 CGAATTACATGATCCCCCATGTTGTGCAAAAAGCGGTAGCTCTTCCGTCTCCATC 3118
Db 604 CGAATTACATGATCCCCCATGTTGTGCAAAAAGCGGTAGCTCTTCCGTCTCCATC 545
Qy 3119 GTTGTCAAGATAGTTGGCCGAGTGTATCATCATCATGTTATGGCAGCATGCTATAT 3178
Db 544 GTTGTCAAGATAGTTGGCCGAGTGTATCATCATCATGTTATGGCAGCATGCTATAT 485
Qy 3179 TCTCTTACGTGATGCAATCCGTAAGATGCTTTCTGTGACTGTGATGACTCAACAG 3238
Db 484 TCTCTTACGTGATGCAATCCGTAAGATGCTTTCTGTGACTGTGATGACTCAACAG 425
Qy 3239 TCATTTGAGAAATAGTATAGCGCGGACCGAGTGTCTTCCCGGCTCAATACGGAT 3298
Db 424 TCATTTGAGAAATAGTATAGCGCGGACCGAGTGTCTTCCCGGCTCAATACGGAT 365
Qy 3299 AATACCGCGCACATAGCAGAACTTTAAAGTGCATCATTTGAAAAAGTCTTCCGGG 3358
Db 364 AATACCGCGCACATAGCAGAACTTTAAAGTGCATCATTTGAAAAAGTCTTCCGGG 305
Qy 3359 CGAAACTCTCAAGATCTTACCGCTGTGAGATCAAGTGTGATACCACTCGTCA 3418
Db 304 CGAAACTCTCAAGATCTTACCGCTGTGAGATCAAGTGTGATACCACTCGTCA 245
Qy 3419 CCCAATCTGATCTTCAAGATCTTTTACTTCAACGCGTTTCTGGGTAGCAAAAACGGA 3478
Db 244 CCCAATCTGATCTTCAAGATCTTTTACTTCAACGCGTTTCTGGGTAGCAAAAACGGA 185
Qy 3479 AGGCAAAATGCGCAAAAAGGAATAAGGCGACAGGAAATGTGGAATCTCATCTC 3538
Db 184 AGGCAAAATGCGCAAAAAGGAATAAGGCGACAGGAAATGTGGAATCTCATCTC 125
Qy 3539 TTCTTTTTCATATTAATGAAGCATTTATCAGGGTATTTGTCTCATGAGCGATACATA 3598
Db 124 TTCTTTTTCATATTAATGAAGCATTTATCAGGGTATTTGTCTCATGAGCGATACATA 65
Qy 3599 TTTGAATGATTTTGAATAAATTAAGAGGTTCCGCGCACATTTCCCGAAAAAGTG 3658
Db 64 TTTGAATGATTTTGAATAAATTAAGAGGTTCCGCGCACATTTCCCGAAAAAGTG 5
Qy 3659 CCAC 3662
Db 4 CCAC 1

```

```

RESULT 3
ASPREN1/c
LOCUS ASPREN1 3626 bp DNA linear SYN 15-MAY-1998

```

```

DEFINITION Artificial sequences, plasmid vector pGreen-1.
ACCESSION Y09373
VERSION Y09373.1 GI:1684624
KEYWORDS blaTEM gene; lacZ-Phoc fusion gene; multiple cloning site; T3 promoter.
SOURCE ORGANISM
  Bacteriophage T3
  Artificial construct
  synthetic construct
  artificial sequence.
REFERENCE
  1 Thaler, M.C., Berlucchi, F., Schippa, S., Selan, L. and Rossolini, G.M.
    Bacterial acid phosphatase gene fusions useful as targets for
    cloning-dependent insertional inactivation
    Biotechnol. Prog. 14 (2), 241-247 (1998)
JOURNAL
  98215688
MEDLINE
  9548775
PUBMED
  2 (bases 1 to 3626)
REFERENCE
  Rossolini, G.M.
  Direct Substitution
  Submitted (11-NOV-1996) G.M. Rossolini, Dip. Biologia
  Molecolare-Sez., Microbiologia, Univ. di Siena, via Laterina N.8,
  1-53100 Siena, ITALY
  Related sequences: X64444 and X52328.
  Location/Qualifiers
    1..3626
      /organism="synthetic construct"
      /mol_type="other DNA"
      /db_xref="taxon:32630"
      /plasmid="pGreen-1"
      129..989
      /gene="blaTEM"
      129..989
      /gene="blaTEM"
      /codon_start=1
      /transl_table=11
      /protein_id="CAA70547.1"
      /db_xref="GI:1684625"
      /translation="MSIQHRYVALIPPEAFCLPYRPAHRETVKVKDAEPLGARVY
      IEIDLSGKILBSFPEREPFMSTFVLCGAVLSRIDGQEOGLRIHYSONDLYE
      YSPVTEKHLTDGNTVRELCSAATMSQNTANLLTTTGGPKELTAFLEHMGDVTFL
      DRWPELNEAIIPNDERDTTPVMAATTLRLTGLLETLASROOLIDMWDKVAQPL
      LRSLPAGWFIADKSGAGERSGIIIALPDGKPSRIVIVYTTGSGATWDERRQIA
      EIASLTKHW"
      2139..2154
      primer_bind
      /note="pUC M13/Reverse Sequencing primer recognition site"
      2146..2934
      gene
      /gene="lacZ-Phoc fusion"
      2146..2934
      CDS
      /gene="lacZ-Phoc fusion"
      /codon_start=1
      /transl_table=11
      /protein_id="CAA70548.1"
      /db_xref="GI:1684626"
      /translation="MTMTIPPSAOLTLTKGNKSSVPAVAALSLVDPAIPAGNDATTK
      PDLYVKNBOAIDISLKLPPPPVGSIFINDAMYEKGMLENTERGKQADADLA
      AGSVATRFSAFEPYITKDSPELYKILTMIDAGDLATRSKAKENMRIRPPRYCT
      ETCNKKQKLTSTNGSPSGHTSIGVATLVLEVPANDALIERVQLGQSRVIG
      YHWQSDVARIVGSAVATLHSDPAFOQLAKAKOEFKSKOK"
      2171..2190
      promoter
      /gene="lacZ-Phoc fusion"
      /note="T3 promoter"
ORIGIN
  Query Match 72.4%; Score 2650.6; DB 12; Length 3626;
  Best Local Similarity 84.0%; Pred. No. 0;
  Matches 3079; Conservative 0; Mismatches 544; Indels 42; Gaps 6;
Qy 1 CTAATTTGTAAGGTTAATTTTGTGTAATTCGCTTAAATTTTGTGTAATCAGCTC 60
Db 3626 CTAATTTGTAAGGTTAATTTTGTGTAATTCGCTTAAATTTTGTGTAATCAGCTC 3567
Qy 61 ATTTTAAACCAATAGCGCAAAATCGGCAAAATCCCTTATTAATCAAAAGATAGACGA 120
Db 3566 ATTTTAAACCAATAGCGCAAAATCGGCAAAATCCCTTATTAATCAAAAGATAGACGA 3507

```

QY 121 GATAGGGTTGAGTGTGTTCCAGTTTGGAACAAAGATCCATTAATTAAAGACGTGACATC 180  
DB 3506 GATAGGGTTGAGTGTGTTCCAGTTTGGAACAAAGATCCATTAATTAAAGACGTGACATC 3447  
QY 181 CAAGCTCAAAAGGCGAAAAACCGTCTATCAGGGCGATGAGCCACTGATGTAACCAATCAC 240  
DB 3446 CAAGCTCAAAAGGCGAAAAACCGTCTATCAGGGCGATGAGCCACTGATGTAACCAATCAC 3387  
QY 241 CTATCAAGTTTGTGGGGTCCAGGTGCGCTTAAGCATTAATCGGAACCTTAAAGGAG 300  
DB 3386 CTATCAAGTTTGTGGGGTCCAGGTGCGCTTAAGCATTAATCGGAACCTTAAAGGAG 3327  
QY 301 CCCCCGATTTAGAGCTTGAACGCGGAAAGCCGCGAATGAGGAGAAAGAAAGAA 360  
DB 3326 CCCCCGATTTAGAGCTTGAACGCGGAAAGCCGCGAATGAGGAGAAAGAAAGAA 3267  
QY 361 AGGAAAGAGAGCGGGGCTTAGGAGCGCTGGCAAGTGAAGCGATGACGCTGCGGTAACAC 420  
DB 3266 AGGAAAGAGAGCGGGGCTTAGGAGCGCTGGCAAGTGAAGCGATGACGCTGCGGTAACAC 3207  
QY 421 CACACCCGCGCGCTTAATGCGCCGCTACAGAGCGCGTCCCATTCGCCATTCAAGCTGCG 480  
DB 3206 CACACCCGCGCGCTTAATGCGCCGCTACAGAGCGCGTCCCATTCGCCATTCAAGCTGCG 3147  
QY 481 CAACGTGTGGGAAGGCGGATCGGTGCGGCGCTTTCGCTATTACGCCAGCTGCGAAAG 540  
DB 3146 CAACGTGTGGGAAGGCGGATCGGTGCGGCGCTTTCGCTATTACGCCAGCTGCGAAAG 3087  
QY 541 GGGATGTGTGTCAGAGCGCATTAAGTGTGGTAACGCGAGGGTTTCCAGTACAGAGCTTG 600  
DB 3086 GGGATGTGTGTCAGAGCGCATTAAGTGTGGTAACGCGAGGGTTTCCAGTACAGAGCTTG 3027  
QY 601 TAAACGACGCGCATGAGCGCGCTCGTTCATTACAGTTTGTGAACCCGAGGAGACG 660  
DB 3026 TAAACGACGCGCATGAGCGCGCTCGTTCATTACAGTTTGTGAACCCGAGGAGACG 2567  
QY 661 GCAGACTCGCGATGCAAAATGTGTTTACAGGTAATGAGACATGAAGTGTGACAC 720  
DB 2966 GCGCCCTCGAGGTCCAGCGTATCGATAGCTTATTTCTGTATTTGTGCAAAATTC 2907  
QY 721 GCTGCAGAACGCGAGCTAGTTAACCTTAGAAAGATATCATATTGTGACGTAACTTAA 780  
DB 2906 TGTGTGCTTTTCGTAACTGCGCTGAATGCGGATCGGAATGATATGTGCGAAGCG 2847  
QY 781 AGATATCATGCGTAATAATTGAGCGCATGAGATCTGTATTCGACTACATATAGGCGAAT 840  
DB 2846 GCTGAAC-----GACATTCGCGCGGCGATCCACATCACTTGCAGTGAATGCGCGA 2794  
QY 841 TGGGTACCGGAGCGCCCTCTGAGAGTGCACGTAATGATTAAGCTTGAATCGAATTCCTGC 900  
DB 2793 AATCACCCGCGCTGTCTCGAGTGAATACCCGTTCCAGATGCGATCTCTGAATTCGCG 2734  
QY 901 AGCCCGGAGGATCCAGTGTCTTAGAGCGCGCCACCGCGGTGAGCTTCAGCTTTTGT 960  
DB 2733 GTTC-----ACTTCGCGACAGCATTAATGCGGTGCCAGCGGATGCGATTTGCGG 2678  
QY 961 TCCCTTTAGTAGAGGTTAAATTAGATCCATGCGTCAATTTTACGAGACTATCTTTAG 1020  
DB 2677 GGTAGATCGTTGTGAGAGATTTTCTGA-TCTTTGTATTTACAGGTTTCTGTGCG 2619  
QY 1021 GGTATATCTAGCGCATCAAGATCATATGTCGGGCTTTTTCGGGCTAGTCATCCG 1080  
DB 2618 TAAAGCGAAAGGACGAGTGCATGTATGTTTGTGGCGAGCGGATGCGAAGTCA 2559  
QY 1081 CAAGCTGGGCTATCTGGCATTCGGGAGGAAAGACCGGTGCTTTTCCGCGAGGTTG 1140  
DB 2558 CCGGCACTCTCATATATGTGTCAAGCTTTATCAAGCTCCGAGAGTCTTTTTCGCTT 2499  
QY 1141 AAGCGCATGGAAGAGTTTGCAGAGTGAATGCTGTGCTGATTCAGCTTGAAGGAAAC 1200  
DB 2498 ATCGATAGCGGAATGCGCTGAATAATCGGTGCGACCCCGCGGCGCAAGGTGCA 2439

QY 1201 GCAGTTTACATGATGATTCGGGAAGTGTGCCATGCA--CGCTTTAAAGGTGAAC 1257  
DB 2438 TTTCCCTGTGCTGTGTTTCCGCGCTCGGTATTCGAGCATACGCGCTTTCATCAT 2379  
QY 1258 TGTTCGTTACGGCCACTTGGATACAGTTCGTGCGGCTTTTCCGGACACAGTTCCGA 1317  
DB 2378 GCGTATCATTTAAACGTAATATGCCAGCTTCCGGCGGTGCGGTAAAGTTCAAG 2319  
QY 1318 TGGTCACCCGAAAGCGCATGACACCCGAAACAATCCGCGACAGCCGGAATCCGCTG 1377  
DB 2318 CTGTGATAGCTGTTCATTTTTCAGATATATTAAT-----CCGCG 2277  
QY 1378 CCGGTGTCAGATTAAATGACAGCGGTGCGCGCTGGGATTAATGCTCAGACGACG 1437  
DB 2276 TTGTGTGTGTCATGTTGCTGCTCCGGATTCGAGATTCATGTTCTTAGAGCCGCGC 2217  
QY 1438 ATCTGCTGATGATCCGCGAATAATGACATGATACCCGTGAGTTAACCCGCGGCGCG 1497  
DB 2216 ACCGCGGTGAGCTTCAGCTTTTGTTCCTTTAGTAGAGGTTAAT-----GCGCG 2166  
QY 1498 CTGTGCGTATCATGATGATGATGCTGTTTCGTGTGAATTGTTATCCGTCACAAATTC 1557  
DB 2165 CTGTGCGTATCATGATGATGATGCTGTTTCGTGTGAATTGTTATCCGTCACAAATTC 2106  
QY 1558 ACACAAATTAAGAGCCGGAAGCATTAAGTGAAGCTGTGGGTGCTTAATGATGAGCTA 1617  
DB 2105 ACACAAATTAAGAGCCGGAAGCATTAAGTGAAGCTGTGGGTGCTTAATGATGAGCTA 2046  
QY 1618 ACTCATTAATTAATGTTGGTGGCTCACTGCGCGCTTTCAGTTCGGAACCTGTGTGCA 1677  
DB 2045 ACTCATTAATTAATGTTGGTGGCTCACTGCGCGCTTTCAGTTCGGAACCTGTGTGCA 1986  
QY 1678 GCTGATTAATGATTCGCGCAACGCGGAGAGCGGTTTGTGTTGGAGGCTCTTC 1737  
DB 1985 GCTGATTAATGATTCGCGCAACGCGGAGAGCGGTTTGTGTTGGAGGCTCTTC 1926  
QY 1738 GCGTTCCTCGCTCATGATCTGCTGCGCTGCTGTGCTGCTGCGGAGCGGTATCAGC 1797  
DB 1925 GCGTTCCTCGCTCATGATCTGCTGCGCTGCTGCTGCTGCTGCGGAGCGGTATCAGC 1866  
QY 1798 TCACTCAAAGCGGTAAATGAGTTATCAAGATCAGGGGATTAAGCGGAAGAAACAT 1857  
DB 1865 TCACTCAAAGCGGTAAATGAGTTATCAAGATCAGGGGATTAAGCGGAAGAAACAT 1806  
QY 1858 GTGAGCAAAAGCGCGCAAAAGCGCAAGAACCGTAAAGCGCGTTCGCGCTTTT 1917  
DB 1805 GTGAGCAAAAGCGCGCAAAAGCGCAAGAACCGTAAAGCGCGTTCGCGCTTTT 1746  
QY 1918 CCATAGGCTCCGCGCCCTGACAGCATCAAAATGAGAGCTGAAGTCAAGAGTGGCG 1977  
DB 1745 CCATAGGCTCCGCGCCCTGACAGCATCAAAATGAGAGCTGAAGTCAAGAGTGGCG 1686  
QY 1978 AAACCCGACAGGCTTAATAAGATCAAGAGCTTTCCCTGGAAGTCCCTGTGCGCTC 2037  
DB 1685 AAACCCGACAGGCTTAATAAGATCAAGAGCTTTCCCTGGAAGTCCCTGTGCGCTC 1626  
QY 2038 TCTGTGTGCAACCTGCGCTTACCGGATCACTGTGCGCTTTCCTTCCGGAAGGCT 2097  
DB 1625 TCTGTGTGCAACCTGCGCTTACCGGATCACTGTGCGCTTTCCTTCCGGAAGGCT 1566  
QY 2098 GGGCTTTCTCATGATGATCAGCTGATGATCTCAATTCGATGATGATGATGATGAT 2157  
DB 1565 GGGCTTTCTCATGATGATCAGCTGATGATCTCAATTCGATGATGATGATGATGAT 1506  
QY 2158 GCTGGCTGTGTCAGAACCCCGGTTCAAGCCGACCGCTGCGCTTATCCGTTAACTA 2217  
DB 1505 GCTGGCTGTGTCAGAACCCCGGTTCAAGCCGACCGCTGCGCTTATCCGTTAACTA 1446  
QY 2218 TGTCTTGAATCAACCGGTAAAGACAGCTTAATGCACTGAGAGAGCACTGATTA 2277  
DB 1445 TGTCTTGAATCAACCGGTAAAGACAGCTTAATGCACTGAGAGAGCACTGATTA 1386  
QY 2278 CAGATTTAGCAGAGGATATGTAAGCGGTGCTACAGAGTTCTTGAAGTGTGCGCTTA 2337

Db 1385 CAGATTAGCAGAGGAGTATGTAGGGGCTCAAGAGTTCTTGAAAGTGTGGCTTAA 1326  
 Qy 2338 CTAGCGCTACACTAGAGAGCAAGTATTTGTATCTGCGCTCTGCTGAAGCCAGTTACCTT 2397  
 Db 1325 CTACGGCTACACTAGAGAGCAAGTATTTGTATCTGCGCTCTGCTGAAGCCAGTTACCTT 1266  
 Qy 2398 CGGAAAAAGATGGTACTCTTGATCCGGCAAAACAAACCCGCTGGTAGGGGTGTTT 2457  
 Db 1285 CGGAAAAAGATGGTACTCTTGATCCGGCAAAACAAACCCGCTGGTAGGGGTGTTT 1206  
 Qy 2458 TTTTGTTCAGCAGCAGATTTACGCGAGAAAAAGATCTCAAGAGATCTTTGAT 2517  
 Db 1205 TTTTGTTCAGCAGCAGATTTACGCGAGAAAAAGATCTCAAGAGATCTTTGAT 1146  
 Qy 2518 CTTTTCAGCGGCTCTGACGCTCAGTGGAAAGAAAACTCAGCTTAAGGATTTTGGTCAT 2577  
 Db 1145 CTTTTCAGCGGCTCTGACGCTCAGTGGAAAGAAAACTCAGCTTAAGGATTTTGGTCAT 1086  
 Qy 2578 GAGATTAATCAAAAAAGATCTTCACTAGATCTTTTAATTAATAATGAAGTTTAAATC 2637  
 Db 1085 GAGATTAATCAAAAAAGATCTTCACTAGATCTTTTAATTAATAATGAAGTTTAAATC 1026  
 Qy 2638 AATCTAAAGTATATATAGATTAACCTTGGCTGACAGTACCAGTACTTAATCAGTGAGAGC 2697  
 Db 1025 AATCTAAAGTATATATAGATTAACCTTGGCTGACAGTACCAGTACTTAATCAGTGAGAGC 966  
 Qy 2698 ACCATCTCAGCGATCTGTCTATTTCTGTTACATCCATAGTGGCTGACTCCCGCTGCTGA 2757  
 Db 965 ACCATCTCAGCGATCTGTCTATTTCTGTTACATCCATAGTGGCTGACTCCCGCTGCTGA 906  
 Qy 2758 GATTAATCAATAGGAGGAGGCTTACCATCTGGCCCCGAGTGTGCAATGATACCGCGAGA 2817  
 Db 905 GATTAATCAATAGGAGGAGGCTTACCATCTGGCCCCGAGTGTGCAATGATACCGCGAGA 846  
 Qy 2818 CCCACGCTCAGCGGCTCCAGATTTATCAGCAATTAACAGCAGCGAGCGAGGCGAGACG 2877  
 Db 845 CCCACGCTCAGCGGCTCCAGATTTATCAGCAATTAACAGCAGCGAGCGAGGCGAGACG 786  
 Qy 2878 CAGAAAGTGTCTGTGCAACTTTATCCGCTCCATCCAGTCTATTAATTTGTCGGGAGAC 2937  
 Db 785 CAGAAAGTGTCTGTGCAACTTTATCCGCTCCATCCAGTCTATTAATTTGTCGGGAGAC 726  
 Qy 2938 TAGAGTAAAGTGTGCGCAAGTTAATAGTTTGGCGCAAGTGTGTCATGCTACAGGCAT 2997  
 Db 725 TAGAGTAAAGTGTGCGCAAGTTAATAGTTTGGCGCAAGTGTGTCATGCTACAGGCAT 666  
 Qy 2998 CGTGTGTCAAGCTCGTGTGTTGGTATGGCTTCAATCAGTCCGGTCCCAACGATCAAG 3057  
 Db 665 CGTGTGTCAAGCTCGTGTGTTGGTATGGCTTCAATCAGTCCGGTCCCAACGATCAAG 606  
 Qy 3058 GCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTAGCTCTTCCGTCCGAT 3117  
 Db 605 GCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTAGCTCTTCCGTCCGAT 546  
 Qy 3118 CGTGTGTCAAGTAAAGTGGCGCGAGTGTATCACTCATGTTATGGCAGCATGTCATTA 3177  
 Db 545 CGTGTGTCAAGTAAAGTGGCGCGAGTGTATCACTCATGTTATGGCAGCATGTCATTA 486  
 Qy 3178 TTCTCTTACTGTCAATGCGCATCCGTAAATGCTTTTCTGTAAGTGTGTAAGTGTGTAAGT 3237  
 Db 485 TTCTCTTACTGTCAATGCGCATCCGTAAATGCTTTTCTGTAAGTGTGTAAGTGTGTAAGT 426  
 Qy 3238 GTCAATGTGAAATAGTGTATGCGCGACCGAGTTGCTCTTCCCGCGCTCAATACGGGA 3297  
 Db 425 GTCAATGTGAAATAGTGTATGCGCGACCGAGTTGCTCTTCCCGCGCTCAATACGGGA 366  
 Qy 3298 TAAATCCGCGCAATAGCAGAACTTTAAAGTGTCTCATCTTGGAAAAAGTTCTTCCGG 3357  
 Db 365 TAAATCCGCGCAATAGCAGAACTTTAAAGTGTCTCATCTTGGAAAAAGTTCTTCCGG 306  
 Qy 3358 GCGAAAACTCTCAAGATCTTAACGCGCTTGAGATCCAGTTCGATGTAAACCATCCGTC 3417

Db 305 GCGAAAACTCTCAAGATCTTACCGCTGTGAGATCCAGTTCGATGTAAACCATCTGTC 246  
 Qy 3418 ACCCACTGATCTTACAGATCTTTTACTTTTACACAGCGCTTTCTGGGTGAGCAAAAAACAG 3477  
 Db 245 ACCCACTGATCTTACAGATCTTTTACTTTTACACAGCGCTTTCTGGGTGAGCAAAAAACAG 186  
 Qy 3478 AAGCAAAATGCCCCCAAAAAAGGAAATTAAGGCGCACAGGAAATGTTAAATTCATCTACT 3537  
 Db 185 AAGCAAAATGCCCCCAAAAAAGGAAATTAAGGCGCACAGGAAATGTTAAATTCATCTACT 126  
 Qy 3538 CTTCCTTTTCAATATTTATTAAGCAATTTATCAGGTTATTTGTCATGACGAGTACAT 3597  
 Db 125 CTTCCTTTTCAATATTTATTAAGCAATTTATCAGGTTATTTGTCATGACGAGTACAT 66  
 Qy 3598 ATTGAATGTATTTAGAAAAATTAACAATAGGAGTTCCGCGCACATTTCCCGAAAAAGT 3657  
 Db 65 ATTGAATGTATTTAGAAAAATTAACAATAGGAGTTCCGCGCACATTTCCCGAAAAAGT 6  
 Qy 3658 GCCAC 3662  
 Db 5 GCCAC 1  
 RESULT 4  
 XXU43957  
 LOCUS XXU43957 3624 bp DNA linear SYN 28-JAN-1999  
 DEFINITION Expression vector pBI07-MRz, complete cds.  
 ACCESSION U43957  
 VERSION U43957.1 GI:4097019  
 KEYWORDS  
 SOURCE Expression vector pBI07-MRz  
 ORGANISM Expression vector pBI07-MRz  
 REFERENCE Ruiz,J., Wu,C.H., Ito,Y. and Wu,G.Y.  
 AUTHORS 1 (bases 1 to 3624)  
 TITLE Direct Submission  
 JOURNAL Submitted (27-DEC-1995) Juan Ruiz, Division of  
 Gastroenterology-Hepatology, University of Connecticut Health  
 Center, Farmington, CT 06030, USA  
 COMMENT In vivo expression vector for the synthesis of a ribozyme against  
 the core region of the Hepatitis B virus pregenomic RNA. It does  
 not recognize the viral core protein but the core region of the  
 viral RNA. pBI07-MRz contains a tandem of ten units of the  
 ribozyme, separated between them by the target sequence recognized  
 by the ribozyme, cloned at the XbaI-HindIII site of pBluescript SK  
 (+).  
 FEATURES  
 source  
 1..3624 location/Qualifiers  
 /organism="Expression vector pBI07-MRz"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:45892"  
 complement(657..1425)  
 /note="multiple cloning site"  
 complement(721..766)  
 /product="hammerhead ribozyme"  
 complement(791..836)  
 /product="hammerhead ribozyme"  
 complement(861..906)  
 /product="hammerhead ribozyme"  
 complement(931..976)  
 /product="hammerhead ribozyme"  
 complement(1001..1046)  
 /product="hammerhead ribozyme"  
 complement(1071..1116)  
 /product="hammerhead ribozyme"  
 complement(1141..1186)  
 /product="hammerhead ribozyme"  
 complement(1211..1256)  
 /product="hammerhead ribozyme"  
 complement(1281..1326)  
 /product="hammerhead ribozyme"  
 complement(1351..1396)  
 /product="hammerhead ribozyme"



Promoter complement (1482. .1604)  
/note="Lac-Z"  
gene complement (2639. .3499)  
/gene="ampR"  
CDS complement (2639. .3499)  
/gene="ampR"  
/function="ampicillin resistance marker"  
/codon\_start=1  
/transl\_table=11  
/product="beta-lactamase"  
/protein\_id="A010137.1"  
/db\_xref="GI:4097020"  
/translation="MSIOHRYVALIPFAFCIPVAPHEITVYKDAEDOGARVGY  
IELDLSGKILSFPEREPFPMSTFKVLLCGAVLSRIDAGROUQRTHISQNDLVE  
YSPVTEKHLTDGWTVEELCSAATMSDNTAANLLITIGFKEITAFIHNQDHVTRL  
DRWEPELNEAIPNDERDTPVPMATTKRLKGLBLTASRQOLIDWEDAKVAGPL  
LRSAIPAGFIADKSGAIGERSGRIITIALGPDGKPSRIVIVITTSQATMBERNQIA  
ETGASLIXIM"

## ORIGIN

Query Match 72.2%; Score 2644.8; DB 12; Length 3624;  
Best Local Similarity 84.3%; Pred. No. 0;  
Matches 3096; Conservative 0; Mismatches 512; Indels 64; Gaps 8;

QY 1 CTAAATTGTACGCTTAATTTTGTAAATTCGCTTAATTTTGTAAATCAGCTC 60  
DB 4 CTAAATTGTACGCTTAATTTTGTAAATTCGCTTAATTTTGTAAATCAGCTC 63  
QY 61 ATTTTAAACCAATAGGCGGAATCGGCAAAATCCCTTAATAATCAAAAGATACCGA 120  
DB 64 ATTTTAAACCAATAGGCGGAATCGGCAAAATCCCTTAATAATCAAAAGATACCGA 123  
QY 121 GATAGGCTGAGTGTGTTTCCAGTTTGAACAAGATCCATTTAAAGAACTGAGCTC 180  
DB 124 GATAGGCTGAGTGTGTTTCCAGTTTGAACAAGATCCATTTAAAGAACTGAGCTC 183  
QY 181 CAACGTCAAAAGGCGCAAAACCGTCTATCAGGCGGAGTCCCATTAAGTAAACCATCAC 240  
DB 184 CAACGTCAAAAGGCGCAAAACCGTCTATCAGGCGGAGTCCCATTAAGTAAACCATCAC 243  
QY 241 CTAAATCAAGTTTGGGCTGAGTGCCTTAAGCACTAAATCGAACCCTAAAGGAG 300  
DB 244 CTAAATCAAGTTTGGGCTGAGTGCCTTAAGCACTAAATCGAACCCTAAAGGAG 303  
QY 301 CCCCAGATTTAGAGCTTGAACGGGGAAGCCGCGCAACGTGGCGAAGAAAGGAGAA 360  
DB 304 CCCCAGATTTAGAGCTTGAACGGGGAAGCCGCGCAACGTGGCGAAGAAAGGAGAA 363  
QY 361 AGCGAAAGAGCGGGGCTAGGCGCTGAGCAAGTGAAGCTGACGCTGCGCTAACAC 420  
DB 364 AGCGAAAGAGCGGGGCTAGGCGCTGAGCAAGTGAAGCTGACGCTGCGCTAACAC 423  
QY 421 CACACCGCGCGCTTAATGCGCGCTACAGGGCGGCTCCCATTTGAGGCTGCG 480  
DB 424 CACACCGCGCGCTTAATGCGCGCTACAGGGCGGCTCCCATTTGAGGCTGCG 483  
QY 481 CAACGTGTGGGAAGGGGATCGGTGCGGCGCTTGTGCTAATTAACGCAAGCTGCGAAAG 540  
DB 484 CAACGTGTGGGAAGGGGATCGGTGCGGCGCTTGTGCTAATTAACGCAAGCTGCGAAAG 543  
QY 541 GGAATGTGTGCAAGCGATTTAAGTTGGGTAACGCCAGGGTTTCCAGTCAAGAGCTTG 600  
DB 544 GGAATGTGTGCAAGCGATTTAAGTTGGGTAACGCCAGGGTTTCCAGTCAAGAGCTTG 603  
QY 601 TAAACGACGCGCACTG-----AGCGCGCTCGTTCACTTACGCTTTTGAACCCGTGGA 654  
DB 604 TAAACGACGCGCACTGTTGTAATGCACTCACTAATAGGCGCAATTTGGGATCGCGGCC 663  
QY 655 GGAAGGAGCACTCGCGGTGCAATGTGTTTAAACGCTGAATGAGCAGATGAAGAGCT 714  
DB 664 CCCCCTCGAGGTGCAAGGATGATAGCTTGAATGATGAGAGAGCTTGGATCCAG 723  
QY 715 CGACACGCTGCAAGACAGCACTAGATTAACCTTGAAGAAATATCATATTTGAGCTA 774

DB 724 CTTAGAGTTTCGCTCAGGACTCATGAGCTCT---GAGCATTTGCTTACAGAG 780  
QY 775 GGTAAAGATATATCATCGTAAATTTGACCATGGGATCTGTAAATACACTCATTAAG 834  
DB 781 CTCTGATTCACAGCTTATAGAGTTTCGCTCCTCAGGACTATTCAGTCTTGAAGTGTCTA 840  
QY 835 GCGAATTTGGATACGGGCCCCCTCGAGGTGACGGTATTCATTAAGTGTATGTAAT 894  
DB 841 GTTCAGAGACTCTGATTCACAGCTTAAAGTTTCCTCTCAGGAG----- 886  
QY 895 TCTGCAAGCCCGGGGATCCACTAGTTTGAAGCGGCCCCACCGGGGTGAGCTTCAGC 954  
DB 887 CTATCAGTCTTACAGATTTGCTAGTTTCAGGAGA-----C 921  
QY 955 TTTTGTTCCTTTAGTGAAGGTTAATTAATCCATGCTCAATTTTAAGCAGATATCT 1014  
DB 922 TCTGATTCACAGCTTAAAGATTTCGCTCAGGACTCATGCTCTAGACTGCTAG 981  
QY 1015 TTTAGAGTTAATCTAGCTGACATCAAGATCATATGCTGGCTTTTTCGGCTAGTC 1074  
DB 982 TTCAGAG-----ACTCTGATTCACAGCTTAAAGTTTCGCT--CTCAGGACTCATCA 1033  
QY 1075 ATGCCCCAAGCTGGCGCTATCTGGGCAATCGGGAAGAAAGCCGCTTTTCCCGCG 1134  
DB 1034 GTCTGAGCATTTGCTAAGTTCAAGAGACTCTGATCCAAAGCTTAAAGTTTCGCTCAC 1093  
QY 1135 AGGTGAAGCGGATGAAAGATTTGGCGAGATGACTGCTGCTCATTAAGCTTGAAGC 1194  
DB 1094 GGAATCATATGCTTACAGATTTGCTAAGTGTGCTGAGAGACT--CTGATTCACAGCTTAAAGT 1152  
QY 1195 GAAAACGACGTTTACATGATGATTCGGAAGGTGCGCATGACAGCCCTTAAACGCTG 1254  
DB 1153 TTGCTCTCAGGACTCATGAGTCTGCTGAGCAATTTCTAAGTTCAGAGACTTGGATCCA 1212  
QY 1255 AACTGTGCTCAGGCTCCTGGATACAGCTTGTGCGGCTTTCCGACACAGTTTC 1314  
DB 1213 GCTTAAAGTTTCGCTCAGGACTCATGAGTCTTAAAGTTTCTAAGTTCAGAGACT 1272  
QY 1315 GGAATGTCAGCCGCAAGGCTATCAGAACCGAACAATAACCGGACAGCGGAACTGCG 1374  
DB 1273 CTGATTCACAGCTTAAAGTTTGTCTCTACAGACTATATAGTCTGAGACTTGTCTAAGT 1332  
QY 1375 GTGCGGTGTCATTAATTAAGACAGCGGTGCGGCTGGATTTACGTCAGCGAAGAG 1434  
DB 1333 TCAGAGACTCTGATTCACAGCTTAAAGTTTGTCTCTACAGACTATATAGTCTGAGACT 1392  
QY 1435 GGTATCTGCTGATCCGCAAGAAATGACATGATACCCCTGAGTTAACCGCGGG- 1493  
DB 1393 ATTGCTTAAGCGCGCGCCACCGCGGTGAGCTTCAGCTTTTGTCTCTTAAAGTGGGT 1452  
QY 1494 -----GCGCTTGGGCTTAATCAATGATGATGCTTTCTGTGTAATTTGATCCGC 1547  
DB 1453 TAATTTGACACTTGGGTATCATGCTCATAGCTGTTTCTGTGTAATTTGATCCGC 1512  
QY 1548 TCACAATTCACCAACATACAGCGCGGAAGATTAAGTGAATAGCTTGGGCTGTAAT 1607  
DB 1513 TCACAATTCACCAACATACAGCGCGGAAGATTAAGTGAATAGCTTGGGCTGTAAT 1572  
QY 1608 GATGAGCTTACATTAATTAATTTGCTGCTCATGCGGCTTTCCAGTGGGAAAC 1667  
DB 1573 GATGAGCTTACATTAATTAATTTGCTGCTCATGCGGCTTTCCAGTGGGAAAC 1632  
QY 1668 TGTGTCAGAGCTGCAATTAATTAATTCGCAACCGCGGGAAGAGCGGTTGCTGTAATG 1727  
DB 1633 TGTGTCAGAGCTGCAATTAATTAATTCGCAACCGCGGGAAGAGCGGTTGCTGTAATG 1692  
QY 1728 GGGGCTTTTCGCTTCTGCTCATCTGACTGCTGCGTGGTTCGCTGCGGAGAG 1787  
DB 1693 GGGGCTTTTCGCTTCTGCTCATCTGACTGCTGCGTGGTTCGCTGCGGAGAG 1752  
QY 1788 CGGTATCAGCTTACCAAGCGGTAATACGTTATTCACAGAAATCAAGGATTAACGAG 1847

Dp	1753	CGGTATCAGCTCACCTCAAAAGCCGGTAATACGGTTATCCACGAATCAGGGATTAACGAC	1812
Qy	1848	GAAAGACATGTGAGCAAAAGGCCAGCAAAAGCCAGGAACCGTAAAAAGGCCGCGTTGC	1907
Dp	1813	GAAAGACATGTGAGCAAAAGGCCAGCAAAAGCCAGGAACCGTAAAAAGGCCGCGTTGC	1872
Qy	1908	TGGCGTTTTTCATATAGGCTCCGCCCCCTTGACGAGCATCAAAAAATCGACGCTCAAGTC	1967
Dp	1873	TGGCGTTTTTCATATAGGCTCCGCCCCCTTGACGAGCATCAAAAAATCGACGCTCAAGTC	1932
Qy	1968	AGAGGTGCGCAAAACCGACAGCACTATATAATATACAGGCGTTTTCCCCCTGGAAGCTCCC	2027
Dp	1933	AGAGGTGCGCAAAACCGACAGCACTATATAATATACAGGCGTTTTCCCCCTGGAAGCTCCC	1992
Qy	2028	TCGTGCGCTCTCTGTGTCGACCCCTGCGCTTACCGATATACCTGTCCGCTTTCTCCCTT	2087
Dp	1993	TCGTGCGCTCTCTGTGTCGACCCCTGCGCTTACCGATATACCTGTCCGCTTTCTCCCTT	2052
Qy	2088	CGGGAAGGATGAGCCGCTTTCTATAGCTACGCTGTAGGTAATCTCAAGTTCCGTGTAGGTGC	2147
Dp	2053	CGGGAAGGATGAGCCGCTTTCTATAGCTACGCTGTAGGTAATCTCAAGTTCCGTGTAGGTGC	2112
Qy	2148	TTCCGCTCCAAAGCTGGGCTGTGTCAGCAACCCCGCTTACGCGGACCGGCTGAGCTTAT	2207
Dp	2113	TTCCGCTCCAAAGCTGGGCTGTGTCAGCAACCCCGCTTACGCGGACCGGCTGAGCTTAT	2172
Qy	2208	CCGGTAACTATCGTCTTGAGTCCAACCCGGTATAGACAGACTTATCCGCACTGACAGCAG	2267
Dp	2173	CCGGTAACTATCGTCTTGAGTCCAACCCGGTATAGACAGACTTATCCGCACTGACAGCAG	2232
Qy	2268	CCACTGTGTAACAGGATTATGACAGACCGAGGTAATGAGGCGGTGCTACAGAGTTCTTGAAGT	2327
Dp	2233	CCACTGTGTAACAGGATTATGACAGACCGAGGTAATGAGGCGGTGCTACAGAGTTCTTGAAGT	2292
Qy	2328	GGTGGACCTAACTAGGCTACACTAGGAAGGACAGATATTGGATATCGGCGCTGCTGTAAGC	2387
Dp	2293	GGTGGACCTAACTAGGCTACACTAGGAAGGACAGATATTGGATATCGGCGCTGCTGTAAGC	2352
Qy	2388	CAGTTACCTTGGAAGAAAGATGGTAGCTCTTGATCCGGCAAACAACACGCGCTGTA	2447
Dp	2353	CAGTTACCTTGGAAGAAAGATGGTAGCTCTTGATCCGGCAAACAACACGCGCTGTA	2412
Qy	2448	GCGGTGTTTTTTTTGTTTGGCAAGCAGCAGATTACGCCACAGAAAAAGATCTCAGAAAG	2507
Dp	2413	GCGGTGTTTTTTTTGTTTGGCAAGCAGCAGATTACGCCACAGAAAAAGATCTCAGAAAG	2472
Qy	2508	ATCCCTTGATCTTTTCTACGCGGGGTCTGACGCTCAGTGGAAACGAAACTCAGCTTAAAGGA	2567
Dp	2473	ATCCCTTGATCTTTTCTACGCGGGGTCTGACGCTCAGTGGAAACGAAACTCAGCTTAAAGGA	2532
Qy	2568	TTTTTGTCATAGATATATCAAAAGAGATCTTCACTAGATCTTTTAAATTTAAAGATGAA	2627
Dp	2533	TTTTTGTCATAGATATATCAAAAGAGATCTTCACTAGATCTTTTAAATTTAAAGATGAA	2592
Qy	2628	GTTTTAAATCAATCTAAAGTATATATAGATTAACCTTGCTGACAGTTTACCAATGCTTAA	2687
Dp	2593	GTTTTAAATCAATCTAAAGTATATATAGATTAACCTTGCTGACAGTTTACCAATGCTTAA	2652
Qy	2688	TCAGTGAAGGACCTATCTCAGGCGATCTGTCTAATTTCCGTATCCATAGTTGGCCGATCC	2747
Dp	2653	TCAGTGAAGGACCTATCTCAGGCGATCTGTCTAATTTCCGTATCCATAGTTGGCCGATCC	2712
Qy	2748	CCGTCGTGTAGATTAATCACTACGATACGGAAGGCTTACCATCTGCGCCACAGTCTCAATGA	2807
Dp	2713	CCGTCGTGTAGATTAATCACTACGATACGGAAGGCTTACCATCTGCGCCACAGTCTCAATGA	2772
Qy	2808	TACCGCGAAGCCACGCTCACCGGCTCCAGATTTATCAGCAATTAACCAAGCCAGCCGAA	2867
Dp	2773	TACCGCGAAGCCACGCTCACCGGCTCCAGATTTATCAGCAATTAACCAAGCCAGCCGAA	2832
Qy	2868	GGGCGAGCGCGAGAAAGGTGCTCTGCAACTTTATCCGCGCTCCATCTCAATCTATTAATGTT	2927
Dp	2833	GGGCGAGCGCGAGAAAGGTGCTCTGCAACTTTATCCGCGCTCCATCTCAATCTATTAATGTT	2892

[illegible]

```
/organism="synthetic construct"  
/mol_type="unassigned DNA"
```

ORIGIN /db\_xref="taxon:32630"

Query Match 72.1%; Score 2639; DB 6; Length 3681;  
Best Local Similarity 83.5%; Pred. No. 0;  
Matches 3075; Conservative 0; Mismatches 580; Indels 28; Gaps 6;

3 AAATGTAGAGCTTAATTTTGTATAATTCGCGTTAAATTTTGTATAATTCAGTCTAT 62  
1 AAATGTAAAGCTTAATTTTGTATAATTCGCGTTAAATTTTGTATAATTCAGTCTAT 60  
63 TTTTAAACCAATAGGCGCAAAATCGGCAAAATCCCTTATTAATCAAAAGATAGACCGA 122  
61 TTTTAAACCAATAGGCGCAAAATCGGCAAAATCCCTTATTAATCAAAAGATAGACCGA 120  
123 TAGGGTGAAGTGTCTTCAAGTTTGGAACAAAGATCCACTATTAAGAACTGAGCTTCA 182  
121 TAGGGTGAAGTGTCTTCAAGTTTGGAACAAAGATCCACTATTAAGAACTGAGCTTCA 180  
183 AGCTCAAAAGGCGAAAAACCGTCTATCAGGGGAGTGGCCCACTAGCGTAACATCACCCT 242  
181 AGCTCAAAAGGCGAAAAACCGTCTATCAGGGGAGTGGCCCACTAGCGTAACATCACCCT 240  
243 AATCAAGTTTGTGGGCTGAGGTGCGGTAAAGCACTAAATCGGAACCTTAAAGGAGCC 302  
241 AATCAAGTTTGTGGGCTGAGGTGCGGTAAAGCACTAAATCGGAACCTTAAAGGAGCC 300  
303 CCCGATTTAGAGCTTGAACGGGGAAGCGCGGAACGTGGGGAAGAAAGAGGAAAG 362  
301 CCCGATTTAGAGCTTGAACGGGGAAGCGCGGAACGTGGGGAAGAAAGAGGAAAG 360  
363 CGAAAGAGCGGGCGCTAGAGCGCTGGCAAGTGAAGCGGTCAAGCTCGCGTTAACCA 422  
361 CGAAAGAGCGGGCGCTAGAGCGCTGGCAAGTGAAGCGGTCAAGCTCGCGTTAACCA 420  
423 CACCGCGCGGCTTAATGCGCGCGCTTACAGGCGCGTCCCATTCGCAATTCAGGCTG 482  
421 CACCGCGCGGCTTAATGCGCGCGCTTACAGGCGCGTCCCATTCGCAATTCAGGCTG 480  
483 ACTGTGGGAAGGCGATCGGTGCGGCGCTTTCGCTATTAACCGCGCTGGCGGAAGGG 542  
481 ACTGTGGGAAGGCGATCGGTGCGGCGCTTTCGCTATTAACCGCGCTGGCGGAAGGG 540  
543 GATGTGCTGCAAGCGCATTAAGTTGGGTAACGCAAGGTTTCCAGTCAAGAGCTTTGA 602  
541 GATGTGCTGCAAGCGCATTAAGTTGGGTAACGCAAGGTTTCCAGTCAAGAGCTTTGA 600  
603 AAACGACGCGCAGTGAAGCGCGCTGTTCACTGTTTGAACCGCTGAGAGAGCGGC 662  
601 AAACGACGCGCAGTGAAGCGCGCTGTTCACTGTTTGAACCGCTGAGAGAGCGGC 660  
663 AGACTGCGGCTGCAAAATGTTTGAACGCGTGAAGAGAGTGAAGATGCTGCAACGC 722  
661 GTGGCGGCGCTGTAAGTGTGCACTGCAATGCAATGCTGCTGATGCTGCAATGTTA 720  
723 TGCAGAACGCGCAGTGAATTAACCTTGAAGATATATATTTGAGCTAGCTTTAAAG 782  
721 AGCAGTATATATACATCCGCTATCGTACGTGAATGAGTCAATGAGTCCCGACACCC 780  
783 AATATCATGCTTAATTAATTAAGCATGAGATCTGTAATAGACCTCACTAATAGGCGA 842  
781 GCCAACACCGCTGAGAGCGCGCTGAGAGGCTGTCTGCTCCGCGCATCCGTTACAG 840  
843 GGTACGAGGCGCGCTCGAGGTGAGCGGTATCGA-----TGAAGTGAATTCGAATTC 896  
841 AGCTGTGACCGTCTCGGAGGCTGAGTGTGAGAGGTTTTCACCGTATACCGAAGC 900  
897 CTGACAGCGCGGGAATCACTAGTCTGAGAGGCGCGCAACGCGGTGAGCTTCCAGCTT 956  
901 CCGAGAGCAGTAAGGTGAGTCTTGTGAGCAATTCGCTTAAAGTAAGTAATTTGCTG 960  
957 TTGTTCCCTTAAAGTGAAGGTTAATTAATCCCAT-----GCTCAATTTTACGCAAC 1009

DB 961 TAAAGTCTACTGTGCGATCAACCGCTTCCAGTAGAGACAGAAAGCAATTTGATGTA 1020  
QY 1010 TATCTTTAGAGTTAATCTAGCTGCATCAGGATCATATCTGCGGCTCTTTTCCGCT 1069  
DB 1021 TTTGAGAGAAAGATGCGAGGAATCAATACATTAAGGTGAACCTTCTTGTGTGCT 1080  
QY 1070 CAGTCATCGCCCAAGCTGCGCTATCTGGGCATCGGGGAGGAAGAACCGGCTTTTC 1129  
DB 1081 TCAGCAATGGTATATCTCATGTTTGAATGGCCCTTGAAGATCGGCGCAAGCTTGATGCC 1140  
QY 1130 CCGCAGGTTGAACGCGCATGGAAGAGTTTGCAGAGATGACTGCTGCTGATTCAGCT 1189  
DB 1141 TGAAGTCTGATCTAAGAGATCCCGACCTCGACCGCATTAATATGTTTCCGTAAT 1200  
QY 1190 TGAAGCAAAACGACGTTTACATGATGATTCGGGAAGGTGTGGCCATGACAGCTTTA 1249  
DB 1201 TCAGCGCTTTCATGATGAGACAGGCGCTT--GAATGTTGACGGGATTAACATATAG 1258  
QY 1250 CCGTGAACCTGTTCTGTCAGGCCACTGGGATACAGTTGCTGCGGCTTTTCCGACACA 1309  
DB 1259 CAATGACGCGACAGCAATTAATCTAAGG--AGCAGAAAGGAGGGTATCCACAAAGT 1315  
QY 1310 GTTCCGGAATGTCAGCCCGAAGCGCATCAGCAACCGAACTAATACCGCGCACG----- 1363  
DB 1316 CCAAGCTACCAATTAACGCAAGCCTCAACGCAAGCAGACGAGAGCGGTGATGACGA 1375  
QY 1364 CCGGAATGCGCGTGCCTGCTGAGATTAATGACAGCGGTGCGGCTGGATATTAAGT 1423  
DB 1376 TCMAAATCTTGTACTCTGTCAGAAATCGAAATCATCTTCGTTAATTCMAAAGCGAG 1435  
QY 1424 CAGCGAGAGCGGTTAT---CTGCTGAGATGCGCAGAAATGACATGATGCCCGTG 1479  
DB 1436 AAGCTGATGAAATTCGACTCGAGGGGGGCGCGTACCAAGCTTTGTTCCCTTTA 1495  
QY 1480 AGTTACCGGCGGGCGCTTGGGCTAATCAATGATCAATGCTGTTTCTGTGTAAATTC 1539  
DB 1496 GTGAGGTTAATTCGAGGTTTGGGCTAATCAATGCTAATGCTGTTTCTGTGTAAATTC 1555  
QY 1540 TTAATCGCTCAATTCACACAACATACGAGCGGGAAGTAATTAAGCTTGGGG 1599  
DB 1556 TTAATCGCTCAATTCACACAACATACGAGCGGGAAGTAATTAAGCTTGGGG 1615  
QY 1600 TGCCTAATGAGTGAAGCTAATCAATTAATGCTGCTGCTCACTGCGCGCTTTTCAATC 1659  
DB 1616 TGCCTAATGAGTGAAGCTAATCAATTAATGCTGCTGCTCACTGCGCGCTTTTCAATC 1675  
QY 1660 GGGAAACCTGTCTGCGCAGCTGCAATTAATGATGAGTGGGCAACGCGGGGAGAGCGGTTT 1719  
DB 1676 GGGAAACCTGTCTGCGCAGCTGCAATTAATGATGAGTGGGCAACGCGGGGAGAGCGGTTT 1735  
QY 1720 GCGTATTTGGGCTCTTCCGCTTCTGCTCACTGACTGCTGCGCTCGGTCTGCGCT 1779  
DB 1736 GCGTATTTGGGCTCTTCCGCTTCTGCTCACTGACTGCTGCGCTCGGTCTGCGCT 1795  
QY 1780 GCGCGAGCGGTTATCAGCTCATCAAAAGGCGGTAATACGTTATTCACAAATAGAGGA 1839  
DB 1796 GCGCGAGCGGTTATCAGCTCATCAAAAGGCGGTAATACGTTATTCACAAATAGAGGA 1855  
QY 1840 TAAAGCAGAAAGAACTGTGAGCAAAAGCCAGCAAAAGCCAGAAACCGTAAAGAGC 1899  
DB 1856 TAAAGCAGAAAGAACTGTGAGCAAAAGCCAGCAAAAGCCAGAAACCGTAAAGAGC 1915  
QY 1900 CCGGTTGCTGCGGTTTTCATAGAGCTCGCGCCCTGAGAGCATCAAAAATGACG 1959  
DB 1916 CCGGTTGCTGCGGTTTTCATAGAGCTCGCGCCCTGAGAGCATCAAAAATGACG 1975  
QY 1960 CTCAAGTCAAGGTGCGAAACCGACAGAGCTATAAAGTACAGAGCTTTCCCTG 2019  
DB 1976 CTCAAGTCAAGGTGCGAAACCGACAGAGCTATAAAGTACAGAGCTTTCCCTG 2035  
QY 2020 AAGTCTCTTGTGCGCTCTCTGTTCCAGCCTGCGCTTACCGGATACCTGTCCGCTT 2079  
DB 2036 AAGTCTCTTGTGCGCTCTCTGTTCCAGCCTGCGCTTACCGGATACCTGTCCGCTT 2095



Qy 2080 TCTCCCTTGGGAAAGCGTGGCTTTCTCATAGCTCAGCGTGAAGTATCTCAGTTCCGT 2139  
 Db 2096 TCTCCCTTGGGAAAGCGTGGCTTTCTCAATGCTCAGCGTGAAGTATCTCAGTTCCGT 2155  
 Qy 2140 GTAGGTCGTTGCTCCTCAAGCTGGGCTGTGTGCACGAAACCCCGCTTCAGCCGACCGCTG 2199  
 Db 2156 GTAGGTCGTTGCTCCTCAAGCTGGGCTGTGTGCACGAAACCCCGCTTCAGCCGACCGCTG 2215  
 Qy 2200 CGCCTTATCCGCTAACTATCGTCTTGAAGTCCAAACCCGGTAAAGACGACTTATGCGCACT 2259  
 Db 2216 CGCCTTATCCGCTAACTATCGTCTTGAAGTCCAAACCCGGTAAAGACGACTTATGCGCACT 2275  
 Qy 2260 GGCAGCGCCACTGCTGTAAACAGATTAGACAGACGAGATGTAGGCGGCTGCTACAGAGTT 2319  
 Db 2276 GGCAGCGCCACTGCTGTAAACAGATTAGACAGACGAGATGTAGGCGGCTGCTACAGAGTT 2335  
 Qy 2320 CTGTAAGTGTGGCTTAAGTACGAGTCACTAGAAAGACAGATTTGGTATCTGCGCTCT 2379  
 Db 2336 CTGTAAGTGTGGCTTAAGTACGAGTCACTAGAAAGACAGATTTGGTATCTGCGCTCT 2395  
 Qy 2380 GCTGAAGCCAGTTACCTTGGGAAAGAGTTGTAGCTCTTGATCCGGCAAAACAAACAC 2439  
 Db 2396 GCTGAAGCCAGTTACCTTGGGAAAGAGTTGTAGCTCTTGATCCGGCAAAACAAACAC 2455  
 Qy 2440 CGCTGTGATCGGCTGTTTTTTTGTGTGCAAGACAGATTTACGCGCAAAAAAAGATC 2499  
 Db 2456 CGCTGTGATCGGCTGTTTTTTTGTGTGCAAGACAGATTTACGCGCAAAAAAAGATC 2515  
 Qy 2500 TCAAGAAGATCCTTTGATCTTTCTAAGGGGCTGACGCTGACGTGAAGCAAAACCTCAG 2559  
 Db 2516 TCAAGAAGATCCTTTGATCTTTCTAAGGGGCTGACGCTGACGTGAAGCAAAACCTCAG 2575  
 Qy 2560 TTAAGGATTTTGTGATGAGATTAATCAAAAGATCTTCACTAGATCCTTTAAATTA 2619  
 Db 2576 TTAAGGATTTTGTGATGAGATTAATCAAAAGATCTTCACTAGATCCTTTAAATTA 2635  
 Qy 2620 AAAATGAAGTTTAAATCAATCTAAAGTATATGAGTAACTTGTGTGACAGTTACCA 2679  
 Db 2636 AAAATGAAGTTTAAATCAATCTAAAGTATATGAGTAACTTGTGTGACAGTTACCA 2695  
 Qy 2680 ATGCTTAATCAGTAGAGGACCTATCTCAGCATCTGTCTATTTGGTTCATCATAGTTGC 2739  
 Db 2696 ATGCTTAATCAGTAGAGGACCTATCTCAGCATCTGTCTATTTGGTTCATCATAGTTGC 2755  
 Qy 2740 CTGACTCCCGCTGTGATGATTAATAAGATACGAGATACGAGGAGCTTACCATCTGCCCCAGTGC 2799  
 Db 2756 CTGACTCCCGCTGTGATGATTAATAAGATACGAGATACGAGGAGCTTACCATCTGCCCCAGTGC 2815  
 Qy 2800 TGCATATGATACCGGAGACCCACGCTCACCGGCTTCAGATTTATCAGCAATAAACAGCC 2859  
 Db 2816 TGCATATGATACCGGAGACCCACGCTCACCGGCTTCAGATTTATCAGCAATAAACAGCC 2875  
 Qy 2860 AGCCGGAAGGGCCGAGCGCAGAAAGTGTCTGCAACTTATCCGCTCATCCAGTCTAT 2919  
 Db 2876 AGCCGGAAGGGCCGAGCGCAGAAAGTGTGTCTGCAACTTATCCGCTCATCCAGTCTAT 2935  
 Qy 2920 TAAATGTTGCGGGAAGCTAGATAGTATTCGCACTTAAATGTTTGCAGCAAGTTGT 2979  
 Db 2936 TAAATGTTGCGGGAAGCTAGATAGTATTCGCACTTAAATGTTTGCAGCAAGTTGT 2995  
 Qy 2980 TGCATATGCTACAGGATCGTGTGTGCACGCTCGTCTTTGGTATGCTTCAATCAGTTC 3039  
 Db 2996 TGCATATGCTACAGGATCGTGTGTGCACGCTCGTCTTTGGTATGCTTCAATCAGTTC 3055  
 Qy 3040 CGGTTCCCAAGCATCAAGGCGAGTTACATGATCCCCATGTTGCAAAAAAGCGGTAG 3099  
 Db 3056 CGGTTCCCAAGCATCAAGGCGAGTTACATGATCCCCATGTTGCAAAAAAGCGGTAG 3115  
 Qy 3100 CTCTTGTGGTCTCTCGATCGTTGTCAAGATAGTTGGCCGAGTGTATCACTCATAGT 3159  
 Db 3116 CTCTTGTGGTCTCTCGATCGTTGTCAAGATAGTTGGCCGAGTGTATCACTCATAGT 3175

Qy 3160 TAGGACAGCACTGCATATTTCTTACTAGTCATGCAATCCGTAAGATGCTTTTCGTGAC 3219  
 Db 3176 TAGGACAGCACTGCATATTTCTTACTAGTCATGCAATCCGTAAGATGCTTTTCGTGAC 3225  
 Qy 3220 TGTGATGACTCAACCAAGTCAATCTGAGATAGTATGATGCGGACGAGCTTCTCTTG 3279  
 Db 3236 TGTGATGACTCAACCAAGTCAATCTGAGATAGTATGATGCGGACGAGCTTCTCTTG 3295  
 Qy 3280 CCGGCGCTCAATACGGAATTAATACCGCGCCACATAGCAAGAACTTTAAAGTGTCAATCAT 3339  
 Db 3296 CCGGCGCTCAATACGGAATTAATACCGCGCCACATAGCAAGAACTTTAAAGTGTCAATCAT 3355  
 Qy 3340 TGGAAAGCTTCTTGGGGGGAAGAACTCTAGAGATCTTACCGGTGTGAGATCCAGTTC 3399  
 Db 3356 TGGAAAGCTTCTTGGGGGGAAGAACTCTAGAGATCTTACCGGTGTGAGATCCAGTTC 3415  
 Qy 3400 GATGTAACCCACTGCTGACACCACTGATCTTTCAGATCTTTTACTTTTCAACGCTTTC 3459  
 Db 3416 GATGTAACCCACTGCTGACACCACTGATCTTTCAGATCTTTTACTTTTCAACGCTTTC 3475  
 Qy 3460 TGGGTGACAAAAACAGAAAGCAAAATGCCGCAAAAAAGGGAATTAAGGCGACACGGA 3519  
 Db 3476 TGGGTGACAAAAACAGAAAGCAAAATGCCGCAAAAAAGGGAATTAAGGCGACACGGA 3535  
 Qy 3520 ATGTTGAATCTCACTCTCTCTTCTTCAATATTAATGAAGCAATTAATCAGGGTTATG 3579  
 Db 3536 ATGTTGAATCTCACTCTCTCTTCTTCAATATTAATGAAGCAATTAATCAGGGTTATG 3595  
 Qy 3580 TCTCATGAGCGGATACATATTTGAATGATTTAGAAAAATTAACAATAGGGGTTCCGCG 3639  
 Db 3596 TCTCATGAGCGGATACATATTTGAATGATTTAGAAAAATTAACAATAGGGGTTCCGCG 3655  
 Qy 3640 CACATTTCCCGGAAAGTGCAC 3662  
 Db 3656 CACATTTCCCGGAAAGTGCAC 3678

RESULT 6  
 LOCUS 115651 3681 bp DNA linear PAT 02-Apr-1996  
 DEFINITION Sequence 6 from patent US 5470573.  
 ACCESSION 115651  
 VERSION 115651.1 GI:1250559  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 REFERENCE 1 (bases 1 to 3681)  
 AUTHORS Lubitz, W. and Szostak, M.P.  
 TITLE Immunogens comprising the non-lytic membrane spanning domain of bacteriophages MS2 or PhiX174  
 JOURNAL Patent: US 5470573-A 6 28-NOV-1995;  
 FEATURES  
 source location/Qualifiers  
 1..3681  
 /organism="unknown"  
 /mol\_type="unassigned DNA"

ORIGIN  
 Query Match 72.1%; Score 2639; DB 6; Length 3681;  
 Best Local Similarity 83.5%; Pred. No. 0;  
 Matches 3075; Conservative 0; Mismatches 580; Indels 28; Gaps 6;

Qy 3 AAATGTGAGCGTTAATATTTTGTAAATTCGCGTTAAATTTTGTGTAATCAGCTCAT 62  
 Db 1 AAATGTGAGCGTTAATATTTTGTAAATTCGCGTTAAATTTTGTGTAATCAGCTCAT 60  
 Qy 63 TTTTAAACAATAGCGGGAATTCGCAAAATCCCTTAAATCAAAAGATAGACGGA 122  
 Db 61 TTTTAAACAATAGCGGGAATTCGCAAAATCCCTTAAATCAAAAGATAGACGGA 120  
 Qy 123 TAGGGTGAAGTGTGTTCCAGTTGGAACAAGAGTCACTATTAAAGAGTGGAATCTCA 182  
 Db 121 TAGGGTGAAGTGTGTTCCAGTTGGAACAAGAGTCACTATTAAAGAGTGGAATCTCA 180

QY 183 AGCTAAAGGGGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTAAACCATCACCTT 242  
| | | | |  
Db 181 AGCTAAAGGGGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTAAACCATCACCTT 240  
| | | | |  
QY 243 AATCAAGTTTTTTTGGGGGTGAGAGTCCCGTAAAGCACTAAATCGGAACCTTAAAGGAGCC 302  
| | | | |  
Db 241 AATCAAGTTTTTTTGGGGGTGAGAGTCCCGTAAAGCACTAAATCGGAACCTTAAAGGAGCC 300  
| | | | |  
QY 303 CCGGATTTTAAAGCTTGAAGGGGAAAAGCCGGCGAAAGTGGCGAAGAAAGGAAAGGAAAG 362  
| | | | |  
Db 301 CCGGATTTTAAAGCTTGAAGGGGAAAAGCCGGCGAAAGTGGCGAAGAAAGGAAAGGAAAG 360  
| | | | |  
QY 363 CGAAAGAGCGGGGCGCTAGGGGCGCTGAGAGTGAAGCGGTACCGCTGCGGTTAACCA 422  
| | | | |  
Db 361 CGAAAGAGCGGGGCGCTAGGGGCGCTGAGAGTGAAGCGGTACCGCTGCGGTTAACCA 420  
| | | | |  
QY 423 CACCCGCGGCTTAATGCGCGCTACAGGGCGGCTCCATTGCGCANTTCAGGCTGGCA 482  
| | | | |  
Db 421 CACCCGCGGCTTAATGCGCGCTACAGGGCGGCTCCATTGCGCANTTCAGGCTGGCA 480  
| | | | |  
QY 483 ACTGTGGGAAGGGCGATCGGGTGGGGCTCTTTGCTATTACGCCAGCTGGCGAAAGGG 542  
| | | | |  
Db 481 ACTGTGGGAAGGGCGATCGGGTGGGGCTCTTTGCTATTACGCCAGCTGGCGAAAGGG 540  
| | | | |  
QY 543 GATGTGCTGAAGGCGATTAGTTAGGTTAAGGTTAAGCGCAAGGGTTTTCCAGTGAAGAGTGT 602  
| | | | |  
Db 541 GATGTGCTGAAGGCGATTAGTTAGGTTAAGGTTAAGCGCAAGGGTTTTCCAGTGAAGAGTGT 600  
| | | | |  
QY 603 AAAGGAGCGGCGATGAGAGCGCGCTGTTCACTTCACTTTTGAACCGGTGAAGGACGGG 662  
| | | | |  
Db 601 AAAGGAGCGGCGATGAGAGCGCGCTGTTCACTTCACTTTTGAACCGGTGAAGGACGGG 660  
| | | | |  
QY 663 AGACTCGGCGGCAATATGTGTTTAAAGCGGTGATGAGAGCAATGAAGATCTGACAGC 722  
| | | | |  
Db 661 GTGGCGGCGGCTGATGATGTGCTGCTGACAGTCAATCTGCTGATGAGCGGATGATTA 720  
| | | | |  
QY 723 TGCAGAACAGCGAGCTAGATTAACTTGAAGAAAGTAACTATTTGTAAGCTTAAAG 782  
| | | | |  
Db 721 AGCGAGATATATACCTCGCTATGCTACGTAAGTGGGTATGAGGTGCGCCCGGACAGCC 780  
| | | | |  
QY 783 ATATCATGGGTAAATTTAGACGATGGGATCTGTAAATCACTCACTAATAGGCGAATTTG 842  
| | | | |  
Db 781 GCCAACACCGGCTGACGCGCTGACGCGGCTTGTCTGCTCCGCGCATCCGCTTACAGACA 840  
| | | | |  
QY 843 GGTACCGGGCCCCCTCGAGGTGACGAGTATCGA-----TAAAGCTAATATCGAATTC 896  
| | | | |  
Db 841 AGCTGTGACCGTCTCCGGGAGCTGCAATGTGTCAAGAGTTTTTCAACGCTATCACCGAAGC 900  
| | | | |  
QY 897 CTGCAAGCCGGGGGATTCATAGTTCTAAGCGCGCGCAACCGCGGTGAGCTCCAGTT 956  
| | | | |  
Db 901 CGGAGGCAATTAAGTCGATGCTTTGTAGCAATTCGTCCCTTAATGAAGCAATTTGCTG 960  
| | | | |  
QY 957 TTGTTCCCTTTAGTGAAGGTTAATTAATCCCAT-----GCGTAAATTTTACGCAAC 1009  
| | | | |  
Db 961 TAAAGTCGTACGTGGGATTCACGCTTCAAGTACGACAGAAAGCAATTTGAATTTGTTAA 1020  
| | | | |  
QY 1010 TATCTTTTGAAGGTTAATCTAGCTGCATCAGATCATATCGTGGGCTTTTTCGCGGT 1069  
| | | | |  
Db 1021 TTTTGAAGAAAGATCGCGAGAAAGTCAATACATTAAGATTGAATTTTCTTTGTTGCT 1080  
| | | | |  
QY 1070 CAGTCATCGCCCAAGTGGCGCTATCTGGGCAATCGGGGAGAAAGAACCCGCTGCTTTTC 1129  
| | | | |  
Db 1081 TCGACATGGGTATCTCATGTTTGAATGGCCCTAAGAGATCCGGCAAGCTTGCATGCG 1140  
| | | | |  
QY 1130 CCGGAGGTTGAAGCGGCAATGMAAGTTTGCAGAGATGAAGTGTGTGATGAGGT 1189  
| | | | |  
Db 1141 TGAAGTGAAGTCTAAGAGATCCCGACGCTGACGCCATTAATATATGTTTTCCGTAAT 1200  
| | | | |  
QY 1190 TGAAGCAAAACGACGTTTACATGATGATTCGGAAGAGTGTGGCCATGACGCTTTAA 1249  
| | | | |  
Db 1201 TGAAGCCTTTCATGATGAGACAGGCGGTTT--GAATGTTGAAGGATGAACATTAAG 1258  
| | | | |

QY 1250 CCGTGAAGTGTGCTTCAAGGCACTGGGATACCAAGTGTGGCGGCTTTCCGACACA 1309  
| | | | |  
Db 1259 CAATGACGGAGCAATTAATCTCAACAG--AGCAGAAAAGCAGGCTATCCACAAGT 1315  
| | | | |  
QY 1310 GTTCCGAGTGTACGCCCGAAGCGCATGACCAACCCGAACAATTAACCGGACAG----- 1363  
| | | | |  
Db 1316 CAGCGTACCATTAAGCGCAAGCCTCAACGACGAGCAGACGAGAGCGGTGAGTGA 1375  
| | | | |  
QY 1364 CCGGAAGTGGCCGGGTGTGAGATTATGACACCGGTGGCGGCTGGGATTTACGT 1423  
| | | | |  
Db 1376 TCCAAAGTTTGTACTGTGAGAAATCGAAATCATCTTGTGTTAATCAAAACGCGAG 1435  
| | | | |  
QY 1424 CAGCGAGACCGGTAT-----CTGGCTGATGCCGAGAAATGACATGATATCCCGG 1479  
| | | | |  
Db 1436 AAGCTGAATGAGAAATTCGACCTCGAGGGGGGCCCGGTATCCAGCTTTGTTCCCTTTA 1495  
| | | | |  
QY 1480 AGTTACCGGCGGGCGCGCTTGGCGGTATCATGTGTATAGCTGTTCTGTGTGAATTG 1539  
| | | | |  
Db 1496 GTGAGGTTAATTCGAGCTTGGCGTATCATGTGTATAGCTGTTCTGTGTGAATTG 1555  
| | | | |  
QY 1540 TTATCCGCTCACAAATTCACACAACTACAGACCGGAAAGCATTAAGTAAAGCCTGGGG 1599  
| | | | |  
Db 1556 TTATCGCTCACAAATTCACACAACTACAGAGCCGAAAGCATTAAGTAAAGCCTGGGG 1615  
| | | | |  
QY 1600 TGCCTATGATGAGCTAACTCAATTAATGCGTTGCGCTCACTGCCGCTTTCAGTTC 1659  
| | | | |  
Db 1616 TGCCTATGATGAGAGTAACTCAATTAATGCGTTGCGCTCACTGCCGCTTTCAGTTC 1675  
| | | | |  
QY 1660 GGGAAACCTGTGTGCGACGCTGATTAATGAATCGGCCAAACCGGGGGAGAGCGGTTT 1719  
| | | | |  
Db 1676 GGGAAACCTGTGTGCGACGCTGATTAATGAATCGGCCAAACCGGGGGAGAGCGGTTT 1735  
| | | | |  
QY 1720 GGTATTTGGGCGCTTTCGCTTCTCGCTCACTGACTCGCTGCGCTGCTGCTT 1779  
| | | | |  
Db 1736 GGTATTTGGGCGCTTTCGCTTCTCGCTCACTGACTCGCTGCGCTGCTGCTT 1795  
| | | | |  
QY 1780 GGGGGAACGGGTATGAGTCACTCAAGCGGTAAATCGTTATTCACAGATCAGGGA 1839  
| | | | |  
Db 1786 GGGGGAACGGGTATGAGTCACTCAAGCGGTAAATCGTTATTCACAGATCAGGGA 1855  
| | | | |  
QY 1840 TAAAGCAAGAAAGAAATGTGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 1899  
| | | | |  
Db 1856 TAAAGCAAGAAAGAAATGTGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 1915  
| | | | |  
QY 1900 GCGGTTGCTGCGGTTTTTCCATAGGCTCCGCCCTTGAACGAGCATCAAAATTCAGC 1959  
| | | | |  
Db 1916 GCGGTTGCTGCGGTTTTTCCATAGGCTCCGCCCTTGAACGAGCATCAAAATTCAGC 1975  
| | | | |  
QY 1960 CTCAGTGAAGAGTGGGAAACCGGACAGACCTTAAGATACCAAGCGTTTTCCCTGG 2019  
| | | | |  
Db 1976 CTCAGTGAAGAGTGGGAAACCGGACAGACCTTAAGATACCAAGCGTTTTCCCTGG 2035  
| | | | |  
QY 2020 AAGCTCCTCGTGGCTCTCCTGTTCCGACCTTGCAGCTTAACGGATACCGTCCGCTT 2079  
| | | | |  
Db 2036 AAGCTCCTCGTGGCTCTCCTGTTCCGACCTTGCAGCTTAACGGATACCGTCCGCTT 2095  
| | | | |  
QY 2080 TCTCCCTTCCGGAAGCGTGGGCTTTCTCATAGCTCAAGCTGTAGTATCTCAAGTTCAGT 2139  
| | | | |  
Db 2096 TCTCCCTTCCGGAAGCGTGGGCTTTCTCATAGCTCAAGCTGTAGTATCTCAAGTTCAGT 2155  
| | | | |  
QY 2140 GTAAGTGTTCGCTCAAGAGTGGGCTGTGACGAAACCCCGGTTACAGCCGACGCTG 2199  
| | | | |  
Db 2156 GTAAGTGTTCGCTCAAGAGTGGGCTGTGACGAAACCCCGGTTACAGCCGCTG 2215  
| | | | |  
QY 2200 CGCTTATCCGGTAACTATCGTCTTGAAGTCAACCCGCTTAAGACAGACTTATCGCACT 2259  
| | | | |  
Db 2216 CGCTTATCCGGTAACTATCGTCTTGAAGTCAACCCGCTTAAGACAGACTTATCGCACT 2275  
| | | | |  
QY 2260 GGCAGAGCAACTGTGTAACGATTTAGACAGAGAGTATGAGCGGTCTACAGAGTT 2319  
| | | | |  
Db 2276 GGCAGAGCAACTGTGTAACGATTTAGACAGAGAGTATGAGCGGTCTACAGAGTT 2335  
| | | | |  
QY 2320 CTGAAGTGTGCGCTTAACCTACAGGCTACACTGAAGAGACAGTATTTGATCTGGGCTCT 2379  
| | | | |

Db 2336 CTTGAAGGCTGGCTTAACCTACGCGCTACCTAGAAAGCAAGTATTTGGTATCTGGCGCTCT 2395  
Qy 2380 GCTGAAGCACTTAACCTTCGGAAGAAAGAGTTGGTAGCTCTTGATCCGGCAAAACCAACAC 2439  
Db 2396 GCTGAAGCACTTAACCTTCGGAAGAAAGAGTTGGTAGCTCTTGATCCGGCAAAACCAACAC 2455  
Qy 2440 CGCTGATAGCGGTGCTTTTGTGTGACAGCAAGTATACGGCGCAAGAAAGAGATC 2499  
Db 2456 CGCTGATAGCGGTGCTTTTGTGTGACAGCAAGTATACGGCGCAAGAAAGAGATC 2515  
Qy 2500 TCAAGAAGATCTTTGATCTTTTCTCAACGGGGCTGACGCTGAGTGAAGCAAGAACTGACG 2559  
Db 2516 TCAAGAAGATCTTTGATCTTTTCTCAACGGGGCTGACGCTGAGTGAAGCAAGAACTGACG 2575  
Qy 2560 TTAAGGAGTTTGGTCATGAGATTAATCAAAAGATCTTCACTAGATCCTTTAAATTA 2619  
Db 2576 TTAAGGAGTTTGGTCATGAGATTAATCAAAAGATCTTCACTAGATCCTTTAAATTA 2635  
Qy 2620 AAAATGAAGTTTAAATCAATCAATGAATATATGAGTAACTTGGTCTGACAGTTACCA 2679  
Db 2636 AAAATGAAGTTTAAATCAATCAATGAATATATGAGTAACTTGGTCTGACAGTTACCA 2695  
Qy 2680 ATGCTTAATCAGTAGAGGCACTTATCTGACGATCTGCTATTTGGTTGATCATAGTTGC 2739  
Db 2696 ATGCTTAATCAGTAGAGGCACTTATCTGACGATCTGCTATTTGGTTGATCATAGTTGC 2755  
Qy 2740 CTGACTCCCGCTGCTGATGATTAAGTACGATAACGGAAGGCTTACCATCTGACCCCAAGTC 2799  
Db 2756 CTGACTCCCGCTGCTGATGATTAAGTACGATAACGGAAGGCTTACCATCTGACCCCAAGTC 2815  
Qy 2800 TGCAATGATACCGGGAAGCCACGCTCACCGGCTCCAGATTTATCAGCAATTAACCAAGCC 2859  
Db 2816 TGCAATGATACCGGGAAGCCACGCTCACCGGCTCCAGATTTATCAGCAATTAACCAAGCC 2875  
Qy 2860 AGCGGGAAGGCGGAGCGCAGAGTGTCTGCAACTTATCCGCTCCATCCAGTCTAT 2919  
Db 2876 AGCGGGAAGGCGGAGCGCAGAGTGTCTGCAACTTATCCGCTCCATCCAGTCTAT 2935  
Qy 2920 TAAATGTTGCGGGAAGCTAGAGTAAGTATCCGCAATTAATAGTTTGCACAAGTTGT 2979  
Db 2936 TAAATGTTGCGGGAAGCTAGAGTAAGTATCCGCAATTAATAGTTTGCACAAGTTGT 2995  
Qy 2980 TGCCATTGCTACAGGCACTGCTGTGTCAACGCTCGTGTGTAATGCTTATTCAGCTC 3039  
Db 2996 TGCCATTGCTACAGGCACTGCTGTGTCAACGCTCGTGTGTAATGCTTATTCAGCTC 3055  
Qy 3040 CGGTTCCCAACGATCAAGGCGAGTTACATGATCCGCCATGTTGCAAAAAAGGGTTAG 3099  
Db 3056 CGGTTCCCAACGATCAAGGCGAGTTACATGATCCGCCATGTTGCAAAAAAGGGTTAG 3115  
Qy 3100 CTCTTCGGCTCTCCGATCGTTGTCAAGAGTAAGTTGGCCGACAGTGTATCATCATAGT 3159  
Db 3116 CTCTTCGGCTCTCCGATCGTTGTCAAGAGTAAGTTGGCCGACAGTGTATCATCATAGT 3175  
Qy 3160 TATGCAAGCACTGATTAATCTCTTACTGTATGCAATCCGTAAAGATCTTTTCTGTAC 3219  
Db 3176 TATGCAAGCACTGATTAATCTCTTACTGTATGCAATCCGTAAAGATCTTTTCTGTAC 3235  
Qy 3220 TGTGAGTACTCAACCAAGTCAATCTGAGATAGTATGAGTATGAGGAGCCGAGTCTCTG 3279  
Db 3236 TGTGAGTACTCAACCAAGTCAATCTGAGATAGTATGAGTATGAGGAGCCGAGTCTCTG 3295  
Qy 3280 CCGGCGCTCAATACGGGATTAATACCGGCGCATAGCAAGAACTTTAAAGTCTCATAT 3339  
Db 3296 CCGGCGCTCAATACGGGATTAATACCGGCGCATAGCAAGAACTTTAAAGTCTCATAT 3355  
Qy 3340 TGGAAAACGTTCTTCGGGGCGAAAACCTCAAGATCTTACCGCTGTGAGATCAAGTTC 3399  
Db 3356 TGGAAAACGTTCTTCGGGGCGAAAACCTCAAGATCTTACCGCTGTGAGATCAAGTTC 3415  
Qy 3400 GATGTAACCACTGCTGACCACTGATCTTCAAGATCTTTTCACTGACAGTTTC 3459

Db 3416 GATGTAACCACTGCTGACCACTGATCTTACGATCTTTTACTTTCACACAGTTTC 3475  
Qy 3460 TGGGTGAGCAAAAAACAGAAAGCAAAATGCCGCAAAAAAGGAAATTAAGGCGACACGAA 3519  
Db 3476 TGGGTGAGCAAAAAACAGAAAGCAAAATGCCGCAAAAAAGGAAATTAAGGCGACACGAA 3535  
Qy 3520 ATGTTGAATCTCACTACTCTCTCTTTTCAATATATTAAGGATTAATGAGGCTTATTTG 3579  
Db 3536 ATGTTGAATCTCACTACTCTCTCTTTTCAATATATTAAGGATTAATGAGGCTTATTTG 3595  
Qy 3580 TCTCATGAGCGGATACATATTTGAATGATTTAGAAAATTAACAAATAGGGGTTCCGCG 3639  
Db 3596 TCTCATGAGCGGATACATATTTGAATGATTTAGAAAATTAACAAATAGGGGTTCCGCG 3655  
Qy 3640 CACATTTCCCGGAAAGTGCCAC 3662  
Db 3656 CACATTTCCCGGAAAGTGCCAC 3678

RESULT 7  
A20700 LOCUS A20700 5314 bp DNA linear PAT 22-AUG-1994  
DEFINITION PMTV1 DNA sequence.  
ACCESSION A20700  
VERSION A20700.1 GI:583402  
KEYWORDS  
SOURCE  
ORGANISM  
synthetic construct  
synthetic construct  
artificial sequences.  
REFERENCE  
1 (bases 1 to 5314)  
AUTHORS  
TITLE  
CARRIER-BOUND RECOMBINANT PROTEIN, PROCESS FOR PRODUCING IT AND ITS  
USE AS AN IMMUNOGEN AND VACCINE  
JOURNAL  
Patent: WO 913155-A 4 05-SEP-1991;  
FEATURES  
source  
1. 5314  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"

## ORIGIN

Query Match 71.6%; Score 2623.6; DB 6; Length 5314;  
Best Local Similarity 83.4%; Pred. No. 0;  
Matches 3059; Conservative 0; Mismatches 579; Indels 28; Gaps 6;  
Qy 20 ATTTTGTAAATTCGCGTTAAATTTTGTAAATCAGTCAATTTTAAACCAATAGGCC 79  
Db 1651 ATTTTGTAAATTCGCGTTAAATTTTGTAAATCAGTCAATTTTAAACCAATAGGCC 1710  
Qy 80 GAAATCGGCAAAATCCCTTATTAATCAAAAGATGACCGAGATAGGGTTGAGTGTGT 139  
Db 1711 GAAATCGGCAAAATCCCTTATTAATCAAAAGATGACCGAGATAGGGTTGAGTGTGT 1770  
Qy 140 CCAAGTTTGAACAAAGATCCACTATTAATGAAGAACTCCCAACGTTAAAGGCGGAAAA 199  
Db 1771 CCAAGTTTGAACAAAGATCCACTATTAATGAAGAACTCCCAACGTTAAAGGCGGAAAA 1830  
Qy 200 ACCGTATCAAGGCGAGTGGCCCACTAGTGAACCATCACTATCAAGTTTGTGGGG 259  
Db 1831 ACCGTATCAAGGCGAGTGGCCCACTAGTGAACCATCACTATCAAGTTTGTGGGG 1890  
Qy 1771 CCAAGTTTGAACAAAGATCCACTATTAATGAAGAACTCCCAACGTTAAAGGCGGAAAA 1830  
Db 1891 TCGAGTGGCGGTAAGCACTTAATCGGAACCTTAAAGGAGCCCGCATTTAGAGTTGA 1950  
Qy 260 TCGAGTGGCGGTAAGCACTTAATCGGAACCTTAAAGGAGCCCGCATTTAGAGTTGA 319  
Db 1891 TCGAGTGGCGGTAAGCACTTAATCGGAACCTTAAAGGAGCCCGCATTTAGAGTTGA 1950  
Qy 320 CCGGGAAGAGCGGGAACGTGGCGAGAAAGGAGGAGAAAGGAGGAGCGGGGCT 379  
Db 1951 CCGGGAAGAGCGGGAACGTGGCGAGAAAGGAGGAGAAAGGAGGAGCGGGGCT 2010  
Qy 380 AGGCGCTGCAAGTGTAGCGGTCAACGCTCGGCTAAACCAACCAACCGCGCGCTTAAT 439  
Db 2011 AGGCGCTGCAAGTGTAGCGGTCAACGCTCGGCTAAACCAACCAACCGCGCGCTTAAT 2070

440 GCGCGCTACAGGGCGGGTCCCATTTGGCCATTAGCGCTGCGCACTGTTGGAGAGCGCA 499  
Db GCGCGCTACAGGGCGGGTCCCATTTGGCCATTAGCGCTGCGCACTGTTGGAGAGCGCA 2130  
QY 500 TCGGTGGGGCTCTTCGTATTACCCAGCTGGCGAAGGGGGAGTGTCTGCAAGGGCA 559  
Db TCGGTGGGGCTCTTCGTATTACCCAGCTGGCGAAGGGGGAGTGTCTGCAAGGGCA 2190  
QY 2131 TTAAAGTTGGGTAACGCCAAGGTTTTTCCGAGTCAAGCGTTGTAACGACGGCCAGTGAA 619  
Db TCGGTGGGGCTCTTCGTATTACCCAGCTGGCGAAGGGGGAGTGTCTGCAAGGGCA 2190  
QY 560 TTAAAGTTGGGTAACGCCAAGGTTTTTCCGAGTCAAGCGTTGTAACGACGGCCAGTGAG 619  
Db TCGGTGGGGCTCTTCGTATTACCCAGCTGGCGAAGGGGGAGTGTCTGCAAGGGCA 2190  
QY 2191 TTAAAGTTGGGTAACGCCAAGGTTTTTCCGAGTCAAGCGTTGTAACGACGGCCAGTGAA 2250  
Db TCGGTGGGGCTCTTCGTATTACCCAGCTGGCGAAGGGGGAGTGTCTGCAAGGGCA 2190  
QY 620 CCGCGCTCTTTCATTCACGTTTTTGAACCCGAGAGAGAGGGGACACTCGCGGTGCAAT 679  
Db TCGGTGGGGCTCTTCGTATTACCCAGCTGGCGAAGGGGGAGTGTCTGCAAGGGCA 2190  
QY 2251 TTGTAAATACGACTCACTAATAGGGCGAATTGAGAGTCAACCGCGGTGGCGGCTTAAT 2310  
Db TCGGTGGGGCTCTTCGTATTACCCAGCTGGCGAAGGGGGAGTGTCTGCAAGGGCA 2190  
QY 680 GTTCTTTACAGCGGTATGAGAGCATGAAGATGTTCGACACGCTGCAAGAACGCGAGCTA 739  
Db TCGGTGGGGCTCTTCGTATTACCCAGCTGGCGAAGGGGGAGTGTCTGCAAGGGCA 2190  
QY 2311 ATGTGTCACCTCTCAGTACATCTGTCTGTATGCCCATATGTTAAACGAGTATATACCTC 2370  
Db TCGGTGGGGCTCTTCGTATTACCCAGCTGGCGAAGGGGGAGTGTCTGCAAGGGCA 2190  
QY 740 GATTAAACCTTAAAGATATATCATATTTGAGGTACGTTTAAAGATATATGCTTAAT 799  
Db TCGGTGGGGCTCTTCGTATTACCCAGCTGGCGAAGGGGGAGTGTCTGCAAGGGCA 2190  
QY 2371 CGCTATTCGCTACGTACTGGGTCAATGCTGCGCCCGGACACCGCGAACCGCTGACG 2430  
Db TCGGTGGGGCTCTTCGTATTACCCAGCTGGCGAAGGGGGAGTGTCTGCAAGGGCA 2190  
QY 800 TGAAGCATGGGATCTGTAAATACGACTCTAATGGGCGAATGGGTACCGGGGCCCCCT 859  
Db TCGGTGGGGCTCTTCGTATTACCCAGCTGGCGAAGGGGGAGTGTCTGCAAGGGCA 2190  
QY 2431 CCGCCTGAGCGGGCTGTCTGCTCCGGGATCGCTTACAGCAAGAGTGTACCGCTTCG 2490  
Db TCGGTGGGGCTCTTCGTATTACCCAGCTGGCGAAGGGGGAGTGTCTGCAAGGGCA 2190  
QY 860 CGAGGTGCACGGTAATCGA-----TAAAGTTGATATCGAATTTCTGCAAGCCCGGGGATC 913  
Db TCGGTGGGGCTCTTCGTATTACCCAGCTGGCGAAGGGGGAGTGTCTGCAAGGGCA 2190  
QY 2491 GGAAGTGCATATGTGTCAGAGGTTTACCGCTATCAACGGAACGGCGAGGAGATTAAGTCT 2550  
Db TCGGTGGGGCTCTTCGTATTACCCAGCTGGCGAAGGGGGAGTGTCTGCAAGGGCA 2190  
QY 914 CACTAGTTCTAAGAGCGGCGCGACCGCGAGTGAAGTCTCAAGCTTTTGTTCCTTAAAGAG 973  
Db TCGGTGGGGCTCTTCGTATTACCCAGCTGGCGAAGGGGGAGTGTCTGCAAGGGCA 2190  
QY 2551 GGATGCTTGTGTAGGAATTTGTCCTTAATGAAGAAATGCTGTAAAGTGTCTACTGTGC 2610  
Db TCGGTGGGGCTCTTCGTATTACCCAGCTGGCGAAGGGGGAGTGTCTGCAAGGGCA 2190  
QY 974 GGTATATTAGATCCCAT-----GCGTCATTTTACGACATATCTTTCTAGGGTTAA 1026  
Db TCGGTGGGGCTCTTCGTATTACCCAGCTGGCGAAGGGGGAGTGTCTGCAAGGGCA 2190  
QY 2611 GGATCAACGCTTTCAGTACGACAGAGAACAAATTAATGTGTAAATTTGAGAGAAAGATCG 2670  
Db TCGGTGGGGCTCTTCGTATTACCCAGCTGGCGAAGGGGGAGTGTCTGCAAGGGCA 2190  
QY 1027 TCTAGCTGCATCAGAGTCAATATCGTCGGGTCTTTTTCCGGCTCAGTCAATGCGCAAGCT 1086  
Db TCGGTGGGGCTCTTCGTATTACCCAGCTGGCGAAGGGGGAGTGTCTGCAAGGGCA 2190  
QY 2671 CGAGGAATCAATACATTAAGAGTGAACCTTTGTGTCTTCGACATGGGTAACTCT 2730  
Db TCGGTGGGGCTCTTCGTATTACCCAGCTGGCGAAGGGGGAGTGTCTGCAAGGGCA 2190  
QY 1087 GAGCGTATCTGGGCAATCGGGAGAGAAAGAGCCGCTCTTTCCCGAGGTGAAGCGG 1146  
Db TCGGTGGGGCTCTTCGTATTACCCAGCTGGCGAAGGGGGAGTGTCTGCAAGGGCA 2190  
QY 2731 CATGTTTGAATGGCCCTAGAGAGATCGGCGCAAGCTTCATGCTCGACTGATTAAG 2790  
Db TCGGTGGGGCTCTTCGTATTACCCAGCTGGCGAAGGGGGAGTGTCTGCAAGGGCA 2190  
QY 1147 CATGAAAGATTTGCGCAGATGACTGCTGCTCATTTGACGTTGAGCGAAACGACAGT 1206  
Db TCGGTGGGGCTCTTCGTATTACCCAGCTGGCGAAGGGGGAGTGTCTGCAAGGGCA 2190  
QY 2791 GGATCCCGAGCGCTGACGCAATTAATATATTTTCCGTAAATTCAGCGCTTTCATGAT 2850  
Db TCGGTGGGGCTCTTCGTATTACCCAGCTGGCGAAGGGGGAGTGTCTGCAAGGGCA 2190  
QY 1207 TTACCATGATCAATTCGGGAAGGTGTGCGCATGACGCGCTTTAAACGTTGAATCTGTCTC 1266  
Db TCGGTGGGGCTCTTCGTATTACCCAGCTGGCGAAGGGGGAGTGTCTGCAAGGGCA 2190  
QY 2851 GAGACAGGCGCTTT--GAAATGTTACCGGATGAACAATAATTAACAATGACGACGACATA 2908  
Db TCGGTGGGGCTCTTCGTATTACCCAGCTGGCGAAGGGGGAGTGTCTGCAAGGGCA 2190  
QY 1267 AGGCGACCTGGGATACCAAGTTGTCGCGGGGCTTTTCCGGAACAAGTTCCGATGTGACGC 1326  
Db TCGGTGGGGCTCTTCGTATTACCCAGCTGGCGAAGGGGGAGTGTCTGCAAGGGCA 2190  
QY 2909 AATCTAACAGS---AGCAGGAAAGCGAGGATATCCACAAGTCCAGCTGATCAATTAAG 2965  
Db TCGGTGGGGCTCTTCGTATTACCCAGCTGGCGAAGGGGGAGTGTCTGCAAGGGCA 2190  
QY 1327 CGAAGCGATACGAAACCCGAACATAACGCGACAG-----CCGGAATCGCGTGGCG 1380  
Db TCGGTGGGGCTCTTCGTATTACCCAGCTGGCGAAGGGGGAGTGTCTGCAAGGGCA 2190  
QY 2966 CAACCTTCACGCGACGACGAGCAGAGCGGTCAGTACATCAAACTTTGTTACTC 3025  
Db TCGGTGGGGCTCTTCGTATTACCCAGCTGGCGAAGGGGGAGTGTCTGCAAGGGCA 2190  
QY 1381 GTGTGCGAATTAATGACAGCGGTGCGCGCTGGGAATTAACGTACGCGAGAGCGGTAT- 1439  
Db TCGGTGGGGCTCTTCGTATTACCCAGCTGGCGAAGGGGGAGTGTCTGCAAGGGCA 2190  
QY 3026 GTCAAGAAATTCGAATATCTTTCGTTAAATCCAAAACGCGACAGAACCTGAAATGAAAT 3085  
Db TCGGTGGGGCTCTTCGTATTACCCAGCTGGCGAAGGGGGAGTGTCTGCAAGGGCA 2190  
QY 1440 ---CTGGCTGATGCGCGAAGAAATGACATGATATCCCGTGAATTAACCGGCGGGCGC 1496  
Db TCGGTGGGGCTCTTCGTATTACCCAGCTGGCGAAGGGGGAGTGTCTGCAAGGGCA 2190  
QY 3086 CGACTTCGAGGGGGGGCGCGTACCGAGCTTTGTTCCTTTAGTGAAGGTTAAATTCGGA 3145  
Db TCGGTGGGGCTCTTCGTATTACCCAGCTGGCGAAGGGGGAGTGTCTGCAAGGGCA 2190  
QY 1497 GCTTGGGCTAATATGCTCATACTCTTTCCTGTGTGAATATGTTATCCGCTACCAATTC 1556

Db	3146	GCTGGGGTATCATGGTCTATACCTGTTTCGTGTGAAATTGTTATCTCCGCTCAAAATTC	1616
Qy	1557	CACAACAATACAGAGCCGAGAACATTAAGTGAAGACCTGGGGTGCCTTAATGAGTAGCT	1616
Db	3206	CACACACATAGGAGCCGGAAGATTAATGTGAAGCTGGGGTGCCTTAATGAGTAGCT	3265
Qy	1617	AACTCANATTAATTTGGTTCCGCTCACTCCCGCTTTCCAGTCGGGAAACCTGTCTGCC	1676
Db	3266	AACTCANATTAATTTGGTTCCGCTCACTCCCGCTTTCCAGTCGGGAAACCTGTCTGCC	3325
Qy	1677	AGCTGCATTATGATTCGGCCCAACCGCGCGGGAGAGGGCGTTTGGTATTTGGGGGCTT	1736
Db	3326	AGCTGCATTATTAATTCGGCCCAACCGCGGGAGAGGGCGTTTGGTATTTGGGGGCTT	3385
Qy	1737	CCGCTTCCCTGCTCACTGACTCGCTGCTGCTGTTCCGCTGCGGACCGGATACAG	1796
Db	3386	CCGCTTCCCTGCTCACTGACTCGCTGCTGCTGTTCCGCTGCGGACCGGATACAG	3445
Qy	1797	CTCACTCAAAAGGCGGTAATAGGGTTATCCAGAAATCAAGGGGATTAACGACGAAAGACA	1856
Db	3446	CTCACTCAAAAGGCGGTAATAGGGTTATCCAGAAATCAAGGGGATTAACGACGAAAGACA	3505
Qy	1857	TGTAGCAAAAGCCAGCAAAAGCCAGAAACCGTAATAAGCCCGCTTCTGTGCGTTT	1916
Db	3506	TGTAGCAAAAGCCAGCAAAAGCCAGAAACCGTAATAAGCCCGCTTCTGTGCGTTT	3565
Qy	1917	TCCATAGGCTCCGCCCCCTGACGAGATCACAAAAATCCAGCTCAATGCAAGGTGGC	1976
Db	3566	TCCATAGGCTCCGCCCCCTGACGAGATCACAAAAATCCAGCTCAATGCAAGGTGGC	3625
Qy	1977	GAACCCGACAGAGACTATTAAGATTAACAGGCGTTTCCCGTGAAGCTCCCTGTGGCT	2036
Db	3626	GAACCCGACAGAGACTATTAAGATTAACAGGCGTTTCCCGTGAAGCTCCCTGTGGCT	3685
Qy	2037	CTCCGTTCCGACCCCTGCGGCTTACCGGATACCTGTCCGCTTCTCCCTTCCGAAAGC	2096
Db	3686	CTCCGTTCCGACCCCTGCGGCTTACCGGATACCTGTCCGCTTCTCCCTTCCGAAAGC	3745
Qy	2097	TGGCGCTTTCATAGCTCAGCGTGAAGTATCCAGTTCCAGTGAAGTGTTCGATCCA	2156
Db	3746	TGGCGCTTTCATAGCTCAGCGTGAAGTATCCAGTTTCAGTTTCGATGAAGTGTTCGATCCA	3805
Qy	2157	AACCTGGGCTGTGTGACGAATCCCGGCTTCAAGCCGACCGCTGCGCTTATCCGATAC	2216
Db	3806	AACCTGGGCTGTGTGACGAATCCCGGCTTCAAGCCGACCGCTGCGCTTATCCGATAC	3865
Qy	2217	ATCGCTTGAATCCAAACCGGTAAGACAGCATTAATGACACATGGAGAGGACACATGTA	2276
Db	3866	ATCGCTTGAATCCAAACCGGTAAGACAGCATTAATGACACATGGAGAGGACACATGTA	3925
Qy	2277	ACAGATTAAGCAGAGCGAGTATGTAGCGGTCTACAGATTTCTGAAGTGTGGCTTA	2336
Db	3926	ACAGATTAAGCAGAGCGAGTATGTAGCGGTCTACAGATTTCTGAAGTGTGGCTTA	3985
Qy	2337	ACTACGGGTACACTGAAGAGCAGATTTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCT	2396
Db	3986	ACTACGGGTACACTGAAGAGCAGATTTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCT	4045
Qy	2397	TCGAAAAAAGTTGTAGCTTGTGATTCGGGCAAAACAAACCAACCGCTGTAGCGGTGT	2456
Db	4046	TCGAAAAAAGTTGTAGCTTGTGATTCGGGCAAAACAAACCAACCGCTGTAGCGGTGT	4105
Qy	2457	TTTTTGTTCGAGCGAGCATTAACCGCGCAAAAAAAGAGATCTCAAGAGATCTTTTGA	2516
Db	4106	TTTTTGTTCGAGCGAGCATTAACCGCGCAAAAAAAGAGATCTCAAGAGATCTTTTGA	4165
Qy	2517	TCCTTTCTACGGGGTCTGACGCTCATGTGAACGAAAACTCAGTTTAAGGATTTTGGTCA	2576
Db	4166	TCCTTTCTACGGGGTCTGACGCTCATGTGAACGAAAACTCAGTTTAAGGATTTTGGTCA	4225
Qy	2577	TGAGATTATCAAAAGATCTTCACTTAGATCCTTTAAATTTAAATGAATTTTAAT	2636

Db 4226 TGAATATATCAAAAAGATCTTCACTAGATCCTTTAAATTAATAAATGAAGTTTAAAT 4285  
Qy 2637 CAATCTAAAGTATATATAGTAAACCTTGGCTGACAGTTACCAATGCTTAAATCAGTGGG 2696  
Db 4286 CAACTCTAAAGTATATATAGTAAACCTTGGCTGACAGTTACCAATGCTTAAATCAGTGGG 4345  
Qy 2697 CACCATCTGAGCATCTGCTATTTGGTTCATCATGTTGGCCGATGCTCCCGCGCTGT 2756  
Db 4346 CACCATCTGAGCATCTGCTATTTGGTTCATCATGTTGGCCGATGCTCCCGCGCTGT 4405  
Qy 2757 AGATTAATCAAGTACGAGGAGGCTTACCATCTGAGGCTGAGTCAATGATACCGAGAG 2816  
Db 4406 AGATTAATCAAGTACGAGGAGGCTTACCATCTGAGGCTGAGTCAATGATACCGAGAG 4465  
Qy 2817 ACCCAAGCTCAACCGGCTTCAGATTTATCAGCAATTAACCAAGCCAGCCGAAAGGCGGAGC 2876  
Db 4466 ACCCAAGCTCAACCGGCTTCAGATTTATCAGCAATTAACCAAGCCAGCCGAAAGGCGGAGC 4525  
Qy 2877 GCAGAAGTGTCTCGCAACTTTATCCGCTCCATCCAGTCTTAAATTTGTTGCGGAGAG 2936  
Db 4526 GCAGAAGTGTCTCGCAACTTTATCCGCTCCATCCAGTCTTAAATTTGTTGCGGAGAG 4585  
Qy 2937 CTAGAGTATAGTATCGCCAGTAAATAGTTTGGCAAGTGTGGCAATGCTACAGGCA 2996  
Db 4586 CTAGAGTATAGTATCGCCAGTAAATAGTTTGGCAAGTGTGGCAATGCTACAGGCA 4645  
Qy 2997 TCGTGTGTCAACGCTCGT 3056  
Db 4646 TCGTGTGTCAACGCTCGT 4705  
Qy 3057 GGCAGATTAATCAAGTATCCCGCAATGTGTGCAAAAAGGCTTAAAGTCTCTCGCTCTCCGA 3116  
Db 4706 GGCAGATTAATCAAGTATCCCGCAATGTGTGCAAAAAGGCTTAAAGTCTCTCGCTCTCCGA 4765  
Qy 3117 TCGTGTGTCAAGTATGTGGCCGAGTGTATCATCTGATGATGATGATGATGATGATGATGAT 3176  
Db 4766 TCGTGTGTCAAGTATGTGGCCGAGTGTATCATCTGATGATGATGATGATGATGATGATGAT 4825  
Qy 3177 ATTCTCTTATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3236  
Db 4826 ATTCTCTTATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4885  
Qy 3237 AGTATCTGTGAATAGTATGTGGGCGAGCCGAGTGTCTTGGCCGCTGATTAATACCGG 3296  
Db 4886 AGTATCTGTGAATAGTATGTGGGCGAGCCGAGTGTCTTGGCCGCTGATTAATACCGG 3356  
Qy 3297 ATAAATACCGGCGCATAGCAGAACTTTAAAGTGTCTCATGTTGAATAAGTCTCTCGG 3356  
Db 4946 ATAAATACCGGCGCATAGCAGAACTTTAAAGTGTCTCATGTTGAATAAGTCTCTCGG 5005  
Qy 3357 GGCAGAAATCTCTCAAGATCTTACCGCTGTGTGAATCCAGTTGATGATTAACCACTCGTG 3416  
Db 5006 GGCAGAAATCTCTCAAGATCTTACCGCTGTGTGAATCCAGTTGATGATTAACCACTCGTG 5065  
Qy 3417 CACCAATCTGATCTTCAAGATCTTCACTTCACTCAAGGTTTGGGAGACAAAACAG 3476  
Db 5066 CACCAATCTGATCTTCAAGATCTTCACTTCACTCAAGGTTTGGGAGACAAAACAG 5125  
Qy 3477 GAAGGCAAAATGCGGCAAAAAGGAATTAAGGCGACACAGGAATGTTGAATCTCATAC 3536  
Db 5126 GAAGGCAAAATGCGGCAAAAAGGAATTAAGGCGACACAGGAATGTTGAATCTCATAC 5185  
Qy 3537 TCTTCTTTTCAATATTTATGAAGATTTATCAAGGTTATGTCTCATGAGCGGATACA 3596  
Db 5186 TCTTCTTTTCAATATTTATGAAGATTTATCAAGGTTATGTCTCATGAGCGGATACA 5245  
Qy 3597 TATTTGAATGATTTAAGAAAATAAACAATTAAGGTTTCCGCGACATTTTCCCGAAGAG 3656  
Db 5246 TATTTGAATGATTTAAGAAAATAAACAATTAAGGTTTCCGCGACATTTTCCCGAAGAG 5305  
Qy 3657 TGCCAC 3662  
Db 5306 TGCCAC 5311

RESULT 8  
LOCUS 115649 115649 bp DNA linear PAT 02-APR-1996  
DEFINITION Sequence 4 from parent US 5470573.  
ACCESSION 115649  
VERSION 115649.1 GI:1250557  
KEYWORDS  
SOURCE  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 5314)  
AUTHORS Lubitz,W. and Sostak,M.P.  
TITLE Immunogens comprising the non-lytic membrane spanning domain of bacteriophages MS2 or PhiX174  
JOURNAL Patent: US 5470573-A 4 28-NOV-1995;  
FEATURES  
LOCATION/Qualifiers  
SOURCE 1..5314  
/organism="unknown"  
/mol\_type="unassigned DNA"  
ORIGIN  
Query Match 71.6%; Score 2623.6; DB 6; Length 5314;  
Best Local Similarity 83.4%; Pred. No. 0;  
Matches 3059; Conservative 0; Mismatches 579; Indels 28; Gaps 6;  
Qy 20 ATTTTGTAAATTCGGGTTAAATTTTGTAAATCAGTCTCATTTTAAACCAATAGGCC 79  
Db 1651 ATTTTGTAAATTCGGGTTAAATTTTGTAAATCAGTCTCATTTTAAACCAATAGGCC 1710  
Qy 80 GAAATCGGCAAAATCCCTTATTAATCAAAAGATAGACCGAGATAGGTTGAGTGTGT 139  
Db 1711 GAAATCGGCAAAATCCCTTATTAATCAAAAGATAGACCGAGATAGGTTGAGTGTGT 1770  
Qy 140 CCAAGTTTGAACAAGATGCTCACTTAAAGAGTGAATCCCAAGTCAAAAGGCGAAAA 199  
Db 1771 CCAAGTTTGAACAAGATGCTCACTTAAAGAGTGAATCCCAAGTCAAAAGGCGAAAA 1830  
Qy 200 ACCGTTATCAAGGCGATGAGCCCACTAGTGAACATCACTTATCAAGTTTGTGGG 259  
Db 1831 ACCGTTATCAAGGCGATGAGCCCACTAGTGAACATCACTTATCAAGTTTGTGGG 1890  
Qy 260 TCGAGGTCCGTTAAAGCACTTAATCGGAACCTTAAAGGAGCCCCGATTTAGCTTGA 319  
Db 1891 TCGAGGTCCGTTAAAGCACTTAATCGGAACCTTAAAGGAGCCCCGATTTAGCTTGA 1950  
Qy 320 CGGGGAAAGCCGGGGAAGTGGGGAAGAAAGGAAGGAAGGAAGGAGGCGGCGCT 379  
Db 1951 CGGGGAAAGCCGGGGAAGTGGGGAAGAAAGGAAGGAAGGAAGGAGGCGGCGCT 2010  
Qy 380 AGGCGCTGAGCAAGTGTAGGCGTCAAGCTGCGCTGAACACACACACCGCGCGCTTAAT 439  
Db 2011 AGGCGCTGAGCAAGTGTAGGCGTCAAGCTGCGCTGAACACACACACCGCGCGCTTAAT 2070  
Qy 440 GCGCGCTTACAGGCGCGCTCCCATTTCCCATTTACAGCTTGGGAAAGGCGGA 499  
Db 2071 GCGCGCTTACAGGCGCGCTCCCATTTCCCATTTACAGCTTGGGAAAGGCGGA 2130  
Qy 500 TCGGTGGGGGCTCTTGGCTTATTAAGGCGAGTGGCGAAAGGGGAGTGTGTCGAAGGCGA 559  
Db 2131 TCGGTGGGGGCTCTTGGCTTATTAAGGCGAGTGGCGAAAGGGGAGTGTGTCGAAGGCGA 2190  
Qy 560 TTAAGTTGGGTAAAGCGAGGAGTTTCCAGTCAAGAGTGTGAAGAGAGGCGAGTGAAG 619  
Db 2191 TTAAGTTGGGTAAAGCGAGGAGTTTCCAGTCAAGAGTGTGAAGAGAGGCGAGTGAAG 2250  
Qy 620 CGCGCTCTGTTCAATTCAGTTTGAACCGGTGAGAGCGGCGAGACTCGCGGTGCAAT 679  
Db 2251 TTGTAAATCAAGCTCATATAGGCGGAATGAGCTTCAACCGCGGTGCGCGCTTAAAT 2310  
Qy 680 GTGTTTAAACGCGATGAGAGCAGATGAAGTGTGACACGCTGAGAAACGACAGCTA 739



```
Db 2311 ATGTGACATCTCAGTACATCTGCTGTATGCCGATAGTTAAGCCAGTATATACATC 2370
Qy 740 GATTAAACCCAGAAAGTATCATATTGTGACGTAAAGATATCATCGTAAAT 799
Db 2371 CGCTATGCTACGTACGTGTCATGCGCCGCGCCGACACCCGCGCAACCCGCGTACG 2430
Qy 800 TGACGATGCGATCTGTAAATACGATCACTATAGGCGAATTGGGATCCGGCCCCCT 859
Db 2431 CGCCCTGACGGGCTGTCTGCTCCCGCATCCGCTTACAGACAAAGCTGTACCTCTCG 2490
Qy 860 CGAGGTCGACGGTATGA-----TAACTTGAATTCGAATTCCTGACCCCGGGAGATC 913
Db 2491 GGAAGTCATGTGTCAAGAGTTTTCACCCGTATCACCGAAACGCCGAGGAGTAAAGTTC 2550
Qy 914 CACTAGTCTAGAGCGCGCCGACACCGCGGTGAGCTCCAGCTTTTGTCCCTTAAGTAG 973
Db 2551 GGAATGTTTGAGCAATTCCTTCCTTAAGTAAAGCAATTCGTAAAGTCTGCTACGTGTC 2610
Qy 974 GGTAAATTAGATCCAT-----GCTCAATTTTACGACAGCTATCTTTCTAGGGTTAA 1026
Db 2611 GGATCACCGCTTCCAGTACGACAGAAAGCAATTGATGTAAATTTTCAAGAAAGATCG 2670
Qy 1027 TCTAGCTGATCAGAGATATATCTCGGATCTTTTTCGCGCTAGTATCGCCCAAGCT 1086
Db 2671 CGAGGAAGTCAATACATTAAGATTGAATCTTTTGTGTCTTCAATGAGTATCTTCA 2730
Qy 1087 GAGCTTATCTGGGATCGGGAGAGAAAGCCCGTCTTTTCCCGAGGTTAAACCG 1146
Db 2731 CATGTTTAAATGGCCCTAGAGGATCCGGCAAGCTTGATGCTGCGAGTGTGACTTAA 2790
Qy 1147 CATGAAAGATTGGCCAGAGATGATGCTGCTGATTAAGCTTGAACGAAACGACGT 1206
Db 2791 GGATCCCCGACGCTGACGCCATTAATATGTTTCCGTAAATTCACGCGCTTCATGAT 2850
Qy 1207 TTACCATGATGATTCGGAGAGTGTGGCATGACAGCGCTTAAACGGTAATCTGTCCTC 1266
Db 2851 GAGACAGGCCGTTT--GAATGTACCGGATGAACATTAATTAAGAAAGACGACGACATA 2908
Qy 1267 AGGCACTCGGGATACCAATTCGTGCGGCTTTTCCGACACAGTTCCGATGGTCAAGC 1326
Db 2909 AACTCAACAGG---AGCAGGAAGACGAGGGTATCCCAAAAGTCCAGCGTACCATTAAGC 2965
Qy 1327 CGAAGCGATCAGCAACCCGAAACAATAACCGGCAAG-----CCGGAATCGCCGTCG 1380
Db 2966 CAAGCTCTACGACGCAACGACGAGACGCGTCAAGTAAATCCAAATTTGTATCTC 3025
Qy 1381 GTGTGCAATTAATGACAGCGGTGCGCGCTGGGATTTACGTCAAGGAGACGAGTAT- 1439
Db 3026 GTCAAAATCGAATCATCTTCGGTTAATTCAAAACGCGCAAAACCTGAATGAGAAATT 3085
Qy 1440 ---CCTGGCTGAGTCCGCGAAGATGACATGATACCCCGTGAATCCCGCGGCGC 1496
Db 3086 CGACTCGAGGGGGGCCCGGTACCCAGCTTTTGTTCCTTTAGTAGGGTTAATTCCGA 3145
Qy 1497 GCTTGGCTAATCATGTCATAGCTGTTTCTGTGAAATTGTTACGCTCACAATTC 1556
Db 3146 GCTTGGGCTAATCATGTCATAGCTGTTTCTGTGAAATTGTTACGCTCACAATTC 3205
Qy 1557 CACACAACATAGAGCGCGAAGCATAAAGTAAAGCTTGGAGTCTTAATGATGAGCT 1616
Db 3206 CACACAACATAGAGCGCGAAGCATAAAGTAAAGCTTGGAGTCTTAATGATGAGCT 3265
Qy 1617 AACTCAATTAATGCTTGGCTGCTCACTGCGCGCTTCCAGTCCGGAACCTGTGTGC 1676
Db 3266 AACTCAATTAATGCTTGGCTGCTCACTGCGCGCTTCCAGTCCGGAACCTGTGTGC 3325
Qy 1677 AGTGTATTAATGATCGGCAACGCGCGGGAGAGCGGTTTGGTATTTGGGCGCTCTT 1736
Db 3326 AGTGTATTAATGATCGGCAACGCGCGGGAGAGCGGTTTGGTATTTGGGCGCTCTT 3385
Qy 1737 CCGCTTCTCGGCTCACTGCTGCGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1796
Db 3386 CCGCTTCTCGGCTCACTGCTGCGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3445

Qy 1797 CTGACTCAAAAGCGGTAAATACGGTATCCACAGATCAGGGGATTAACGAGAAAGACA 1856
Db 3446 CTCACTAAAGGCGGTAAATACGGTATCCACAGATCAGGGGATTAACGAGAAAGACA 3505
Qy 1857 TGTGAGCAAAAGCCAGCAAAAGCGAGAAACCGTAAAGAAAGCGCGGTGCTGCGCTTT 1916
Db 3506 TGTGAGCAAAAGCCAGCAAAAGCGAGAAACCGTAAAGAAAGCGCGGTGCTGCGCTTT 3565
Qy 1917 TCCATAGGCTCCCGCCCCCTGACAGCATCAAAAAATGACGCTCAAGTCAAGAGTGC 1976
Db 3566 TCAATAGGCTCCCGCCCCCTGACAGCATCAAAAAATGACGCTCAAGTCAAGAGTGC 3625
Qy 1977 GAAACCGACAGGACTATTAAGATACAGGCGTTTCCCGTGAAGTCCCTCGTGCCT 2036
Db 3626 GAAACCGACAGGACTATTAAGATACAGGCGTTTCCCGTGAAGTCCCTCGTGCCT 3685
Qy 2037 CTCTGTTCCGACCTCGCGCTTAACCGATACCTGTCCGCTTTCCTTCGGGAGAG 2096
Db 3686 CTCTGTTCCGACCTCGCGCTTAACCGATACCTGTCCGCTTTCCTTCGGGAGAG 3745
Qy 2097 TGGCGCTTCTGATAGCTGAGCTGATGATCTCAATTCGCTAGTACGCTGCTCA 2156
Db 3746 TGGCGCTTCTGATAGCTGAGCTGATGATCTCAATTCGCTAGTACGCTGCTCA 3805
Qy 2157 AGCTGGCTGTGTGACGAACCCCGCTTACAGCCGACCGCTGCGCTTATCCGGTAACT 2216
Db 3806 AGCTGGCTGTGTGACGAACCCCGCTTACAGCCGACCGCTGCGCTTATCCGGTAACT 3865
Qy 2217 ATGCTTGAATCCCAACCGGTAAAGACAGCTTATCGCCACTGGACAGACCACTGTA 2276
Db 3866 ATGCTTGAATCCCAACCGGTAAAGACAGCTTATCGCCACTGGACAGACCACTGTA 3925
Qy 2277 ACAGATTAAGCAGACGAGTATGAGCGGTGCTACAGAGTCTTGAAGTGGGCTA 2336
Db 3926 ACAGATTAAGCAGACGAGTATGAGCGGTGCTACAGAGTCTTGAAGTGGGCTA 3985
Qy 2337 ACTAGGCTTACATTAAGAGACAGTATTTGGTATCTGCGCTCTGTAAGCCAGTTACT 2396
Db 3986 ACTAGGCTTACATTAAGAGAGCAGTATTTGGTATCTGCGCTCTGTAAGCCAGTTACT 4045
Qy 2397 TCGAAAAAAGATTGATAGCTCTTGATCCGCAAAACAACCGCTGTAGCGGTGTT 2456
Db 4046 TCGAAAAAAGATTGATAGCTCTTGATCCGCAAAACAACCGCTGTAGCGGTGTT 4105
Qy 2457 TTTTGTGTGACAGCAGATTAACGGCAGAAAAAAGGATCTCAAGAGATCTTTGA 2516
Db 4106 TTTTGTGTGACAGCAGATTAACGGCAGAAAAAAGGATCTCAAGAGATCTTTGA 4165
Qy 2517 TCTTTTCAAGGGGTCTGAGGCTCAGTGAACGAAAACTACGTTAAGGATTTTGCTCA 2576
Db 4166 TCTTTTCAAGGGGTCTGAGGCTCAGTGAACGAAAACTACGTTAAGGATTTTGCTCA 4225
Qy 2577 TGAATTAATCAAAAGGATCTTCACTAGATCTTTTAAATTAATGAAGTTTAAAT 2636
Db 4226 TGAATTAATCAAAAGGATCTTCACTAGATCTTTTAAATTAATGAAGTTTAAAT 4285
Qy 2637 CAATCTAAATTAATGAATTAATCTTGTGACAGTTACCAATCTTAAATCAAGTGA 2696
Db 4286 CAATCTAAATTAATGAATTAATCTTGTGACAGTTACCAATCTTAAATCAAGTGA 4345
Qy 2697 CACTATCTGAGGATCTGCTATTTCTGTCATCCATAGTGTGACCTCCCGTGTGT 2756
Db 4346 CACTATCTGAGGATCTGCTATTTCTGTCATCCATAGTGTGACCTCCCGTGTGT 4405
Qy 2757 AGATTAATAGATAGGAGGGCTTAACATCTGCCCCAGTGTCTCAATGATACCGCAG 2816
Db 4406 AGATTAATAGATAGGAGGGCTTAACATCTGCCCCAGTGTCTCAATGATACCGCAG 4465
Qy 2817 ACCCAGGCTCACCGGCTCCAGATTTATCAGCAATTAACAGCAGCCGGAAGGCGCAGC 2876
Db 4466 ACCCAGGCTCACCGGCTCCAGATTTATCAGCAATTAACAGCAGCCGGAAGGCGCAGC 4525
```

QY 2877 GCAGAGAGTCTCGCACTTTATCCGCTCCATCCAGTCTATTATTTGTTGGCGGAG 2936  
DB 4526 GCAGAGAGTCTCGCACTTTATCCGCTCCATCCAGTCTATTATTTGTTGGCGGAG 4585  
QY 2937 CTAGAGTAGTATTCGCGCACTTATAGTTTGGCGAGTGTGGCTATGCTACAGCA 2996  
DB 4586 CTAGAGTAGTATTCGCGCACTTATAGTTTGGCGAGTGTGGCTATGCTACAGCA 4645  
QY 2997 TCGTGTGTCAGAGTATTCGCGCACTTATAGTTTGGCGAGTGTGGCTATGCTAC 3056  
DB 4646 TCGTGTGTCAGAGTATTCGCGCACTTATAGTTTGGCGAGTGTGGCTATGCTAC 4705  
QY 3057 GCGAGTAGTATTCGCGCACTTATAGTTTGGCGAGTGTGGCTATGCTACAGCA 3116  
DB 4706 GCGAGTAGTATTCGCGCACTTATAGTTTGGCGAGTGTGGCTATGCTACAGCA 4765  
QY 3117 TCGTGTGTCAGAGTATTCGCGCACTTATAGTTTGGCGAGTGTGGCTATGCTAC 3176  
DB 4766 TCGTGTGTCAGAGTATTCGCGCACTTATAGTTTGGCGAGTGTGGCTATGCTAC 4825  
QY 3177 ATTCTCTTACTGTCTATGCTATCCGCTATGCTATGCTATGCTATGCTATGCTATG 3236  
DB 4826 ATTCTCTTACTGTCTATGCTATCCGCTATGCTATGCTATGCTATGCTATGCTATG 4885  
QY 3237 AGTCACTCTGAGATAGTATGCTATGCTATGCTATGCTATGCTATGCTATGCTATG 3296  
DB 4886 AGTCACTCTGAGATAGTATGCTATGCTATGCTATGCTATGCTATGCTATGCTATG 4945  
QY 3297 ATATATACCGCGCCCATAGCAGAACTTTAAAGTGTCTATGCTATGCTATGCTATG 3356  
DB 4946 ATATATACCGCGCCCATAGCAGAACTTTAAAGTGTCTATGCTATGCTATGCTATG 5005  
QY 3357 GCGGAAACTCTCAAGATCTTACCGCTGTGAGATCCAGTTCAGTTCAGTTCAGTTC 3416  
DB 5006 GCGGAAACTCTCAAGATCTTACCGCTGTGAGATCCAGTTCAGTTCAGTTCAGTTC 5065  
QY 3417 CACCAACTGATCTTACGATCTTTACTTACAGAGTTCGCGGAGCAAAAGAG 3476  
DB 5066 CACCAACTGATCTTACGATCTTTACTTACAGAGTTCGCGGAGCAAAAGAG 5125  
QY 3477 GAAAGCAAAATGCGGCAAAAGGAAATGAGGCGACAGGAAATGTTGAATACATAC 3536  
DB 5126 GAAAGCAAAATGCGGCAAAAGGAAATGAGGCGACAGGAAATGTTGAATACATAC 5185  
QY 3537 TCTTCTTTTCAATATATTAAGCATTTATCAGGTTATTTCTCATGAGCGGATCA 3596  
DB 5186 TCTTCTTTTCAATATATTAAGCATTTATCAGGTTATTTCTCATGAGCGGATCA 5245  
QY 3597 TATTGATGTATTGAAATTAAGCAATTAAGGAGTTCGCGGACATTTCCCGAAAG 3656  
DB 5246 TATTGATGTATTGAAATTAAGCAATTAAGGAGTTCGCGGACATTTCCCGAAAG 5305  
QY 3657 TGCCAC 3662  
DB 5306 TGCCAC 5311

RESULT 9  
AF171227/c 3552 bp DNA circular SYN 04-FEB-2000  
LOCUS AF171227 Shuttle vector pSLIRES11, complete sequence.  
DEFINITION AF171227  
ACCESSION AF171227  
VERSION AF171227.1 GI:5764544  
KEYWORDS  
SOURCE Shuttle vector pSLIRES11  
ORGANISM Shuttle vector pSLIRES11  
REFERENCE 1 (bases 1 to 3552)  
AUTHORS Chen, C.-M., Smith, D.M., Peters, M.A., Samson, M.E., Zit, J.,  
Tabin, C.J., and Cepko, C.L.  
TITLE Production and design of more effective avian  
replication-incompetent retroviral vectors  
JOURNAL Dev. Biol. 214 (2), 370-384 (1999)

MEDLINE 99456779  
PUBMED 10525341  
REFERENCE 2 (bases 1 to 3552)  
AUTHORS Chen, C.-M., Samson, M.E.S. and Cepko, C.L.  
TITLE Direct Submission  
JOURNAL Submitted (23-JUL-1999) Genetics, Harvard Medical School, 200  
Longwood Avenue, Boston, MA 02115, USA  
FEATURES  
source location/qualifiers  
1..3552  
/organism="Shuttle vector pSLIRES11"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:101936"  
/lab\_host="Escherichia coli"  
join(2864..3552,1..2208)  
/note="similar to pBluescript"  
2236..2789  
/note="IRES of EMCV"  
2790..2849  
misc\_feature  
misc\_feature  
misc\_feature  
/note="multiple cloning site"  
ORIGIN  
Query Match 71.3%; Score 2610.8; DB 12; Length 3552;  
Best Local Similarity 83.8%; Pred. No. 0;  
Matches 3070; Conservative 0; Mismatches 482; Indels 110; Gaps 6;  
QY 1 CTAAATGTAGAGCTTATATTTGTTTAAATTCGCTTAAATTTGTTTAAATCAAGTC 60  
DB 3552 CTAAATGTAGAGCTTATATTTGTTTAAATTCGCTTAAATTTGTTTAAATCAAGTC 3493  
QY 61 ATTTTAAACCAATAGGCGGAAATCGGCAAAATCCCTTATTAATCAAAAGATAGACGA 120  
DB 3492 ATTTTAAACCAATAGGCGGAAATCGGCAAAATCCCTTATTAATCAAAAGATAGACGA 3433  
QY 121 GATAGGTTGAGTGTGTTCCAGTTTGAACAGAGTCCACTATTAAGAAAGTGAAGTC 180  
DB 3432 GATAGGTTGAGTGTGTTCCAGTTTGAACAGAGTCCACTATTAAGAAAGTGAAGTC 3373  
QY 181 CAAGCTCAAGAGGCGGAAACCGCTTATCAGGCGGATGCGCCACTAGTGAACATCAC 240  
DB 3372 CAAGCTCAAGAGGCGGAAACCGCTTATCAGGCGGATGCGCCACTAGTGAACATCAC 3313  
QY 241 CTAAATCAAGTTTGTGGGTTGAGGTGCGGTAAGCACTAAATCGAACCCTTAAGGAG 300  
DB 3312 CTAAATCAAGTTTGTGGGTTGAGGTGCGGTAAGCACTAAATCGAACCCTTAAGGAG 3253  
QY 301 CCCCCGATTAAGACTTGAAGGCGGAAAGCCGCGCAAGTGGCGAAGAAAGGAGAA 360  
DB 3252 CCCCCGATTAAGACTTGAAGGCGGAAAGCCGCGCAAGTGGCGAAGAAAGGAGAA 3193  
QY 361 AGCGAAAGAGCGGCGCTAGGCGCTGCGCAAGTGTAGCGGTCAAGTGTAGCGGATCA 420  
DB 3192 AGCGAAAGAGCGGCGCTAGGCGCTGCGCAAGTGTAGCGGTCAAGTGTAGCGGATCA 3133  
QY 421 CACACCGCGCGCTTATAGCGCGCTTACAGGCGCGTCCCATTTGGCATTCAAGGCTG 480  
DB 3132 CACACCGCGCGCTTATAGCGCGCTTACAGGCGCGTCCCATTTGGCATTCAAGGCTG 3073  
QY 481 CAATGTTGGAGAGGCGATCGTGGCGGCTCTTCTGCTATTAAGCGCAAGTGGCGAAAG 540  
DB 3072 CAATGTTGGAGAGGCGATCGTGGCGGCTCTTCTGCTATTAAGCGCAAGTGGCGAAAG 3013  
QY 541 GGGATGTGTCAGAGGCGATTAAGTTGGTTAAGCCAGGAGTTTCCAGTCAAGAGCTTG 600  
DB 3012 GGGATGTGTCAGAGGCGATTAAGTTGGTTAAGCCAGGAGTTTCCAGTCAAGAGCTTG 2953  
QY 601 TAAACGAGCGCGAGTGAAGCGCGCTGTTCAATCAAGTTTGAACCCGAGGAGAGCG 660  
DB 2952 TAAACGAGCGCGAGTGAAGCGCGCTGTTCAATCAAGTTTGAACCCGAGGAGAGCG 2913  
QY 661 GCAGACTCGCGGTCAATGTGTTTACAGGTGATGAGAGCAATGAGATGCTGACAC 720  
DB 2912 GCAGACTCGCGGTCAATGTGTTTACAGGTGATGAGAGCAATGAGATGCTGACAC 2853

QY 721 GCTGAGAAACGACACTAGATTAAACCTTAGAAAGATAATCATATTGACGTAGTTAA 780  
Db 2852 GAGGCGG-----CCGACATAGATGATATGATATCCCGGGGCGAGCTCCAAAT 2808  
QY 781 AGATTATCATGCGTAAATTTGACGATGGAGATCTGTATATAGACTCACTATAGGGCGAAT 840  
Db 2807 TCGTATATCATGCGCATGTGATATATCATGTGTTTTTCAAAAGAAACACAGTCCCGGTG 2748  
QY 841 TGGGTACCGGGGCCCCCTCCAGGTCGACGCGTATCGATAGCTTGATATGAAATTCCTGC 900  
Db 2747 TTCGGGGGGGCTTAGCGTTTTTTTAACTCTGACTAAACACAGTAAACATGTGCACCGAG 2688  
QY 901 AGCCCGGGGAGTCCACTAGTTCTAGAGCGGCGGCGACCGGGGTGAGCTCCAGCTTTGT 960  
Db 2687 GCCCGAGATCAGATCCCATACAAATGGGTATCTTTGGGATCTTTGAGCCCTTGTTGA 2628  
QY 961 TCCCTTAGTGAAGGTTAATTAATGCCATGCGTCAATTTTAGCAGACTATCTTTCTAG 1020  
Db 2627 ATAGCTTGAAGAGAGCCATTGTACTCTTCCAACTATCCAACTCAACAGGTGCACT 2568  
QY 1021 GGTAAATCTAGCTGATCAGGATCATATGTCGGGTCTTTTTCGGCTCAGTCAATCGCC 1080  
Db 2567 GG-----GGTTGCGCCCTTTGCAAGGTATCTTATACACGTGCGCTTTTGCC 2519  
QY 1081 CAAGCTGGCGTATCTGGGCAATCGGAGAGAAAGCCCGTGCCTTTTCCGCGAGGTG 1140  
Db 2518 GCAAGGCACTGTGCGCAGGTGGGGGTTCCGCTGCTGCAAAAGGTCTCTACAGACT 2459  
QY 1141 AAGCGCATGGAAGAAGTTTGCAGAGATGACGTGCTGCTGATTAAGCTTAGCGAAAC 1200  
Db 2458 TGTTTGCTTCAAGAAAGCTTTCAGAGG-----ACTGCTCTTACACAAATTCACAGCA 2404  
QY 1201 GCAGCTTTCACATGATGATTCGGGAAAGGTGGCCATGCAAGCTTTTAAACGTGAACTGT 1260  
Db 2403 CATTGATTCCTTTGGCGAAGAGGAGAAAGACCCCTAGGAATGCTGTCAAAAGACAGCG 2244  
QY 1261 TCGTTCAAGGCACTGGGATACAGATTGTCGGGCTTTTCCGACACAGTTCCGATG 1320  
Db 2343 CAGGTTTTCCGGGCTCTACATGTCAAAGACGCAATATGTGAGAAATTAATCATATAG 2284  
QY 1321 TCAGCCCGAAGGCAATCGAACCCGAACAAATACCGGCAAGCCGAGACCTGCGTCCG 1380  
Db 2283 ACAAACGCA-----CACCGCCCTTATTCACAGCGGCTTCGGCAGTAACTGTTAGGG 2231  
QY 1381 GTGTGCAATTAATGACAGCGGTGCGGCTGGGATTAATGTCAGAGAGAGCGGTATC 1440  
Db 2230 GGGGAGGGAGTCTAGAAATGATGATGCTCAAGCTTTTGTCTCTTAAATGAGAGGTAAAT - 2172  
QY 1441 CTGGCTGATGCCCAAAATGACATGATACCCCGTAGTTAACCGGCGGGCGGCTT 1500  
Db 2171 -----TTCCAGCTT 2163  
QY 1501 GGGGTAAATCATGATAGCTGTTTTCTGTGTAATTTGTATCCGCTCACAATTCACA 1560  
Db 2162 GGGCTAATCATAGTCATAGCTGTTTTCTGTGTAATTTGTATCCGCTCACAATTCACA 2103  
QY 1561 CAACATACGAGCCGGAAGCAATAAGTGAAGCTGGGGTGCCTAATGATGATGAGCTAACT 1620  
Db 2102 CAACATACGAGCCGGAAGCAATAAGTGAAGCTGGGGTGCCTAATGATGATGAGCTAACT 2043  
QY 1621 CAATTAATTCGTTGCGCTCACTGCGGCTTTCAGATCGGAAACCTGTGTCAGACT 1680  
Db 2042 CAATTAATTCGTTGCGCTCACTGCGGCTTTCAGATCGGAAACCTGTGTCAGACT 1983  
QY 1681 GCATTAAATATCGGCAACGCGCGGGAGAGAGCGGTGTGGTATTTGGGAGCTCTTCCG 1740  
Db 1982 GCATTAAATATCGGCAACGCGCGGGAGAGAGCGGTGTGGTATTTGGGAGCTCTTCCG 1923  
QY 1741 TTCTCTGCTCACTGACTGCTGCGCTGCGTTCGTTGCGTGGAGAGCGGTATCACTCA 1800  
Db 1922 TTCTCTGCTCACTGACTGCTGCGTTCGTTGCGTGGAGAGCGGTATCACTCA 1863  
QY 1801 CTCAAAAGGCGTAAATCGGTTATTCACAGATCAAGGAGATTAACGAGAAAGAAATGTG 1860

Db 1862 CTCAAAAGGCGTAAATCGGTTATTCACAGATCAAGGAGATTAACGAGAAAGAAATGTG 1803  
QY 1861 AGCAAAAGCCAGCAAAAGGCGAGAAACGTAAAGAGCCGCTGTGCGCTTTTCA 1920  
Db 1802 AGCAAAAGCCAGCAAAAGGCGAGAAACGTAAAGAGCCGCTGTGCGCTTTTCA 1743  
QY 1921 TAGGCTCGGCCCCCTGACAGATCACAAAATTCACAGCTCAAGTCAAGGTGCGAAA 1980  
Db 1742 TAGGCTCGGCCCCCTGACAGATCACAAAATTCACAGCTCAAGTCAAGGTGCGAAA 1683  
QY 1981 CCCGACGAGCTTAAAGATACAGAGCTTTCCCTGGAAGCTCCGCTGTGCGCTCC 2040  
Db 1682 CCCGACGAGCTTAAAGATACAGAGCTTTCCCTGGAAGCTCCGCTGTGCGCTCC 1623  
QY 2041 TGTTCGACCCCTGCTACCGGATACCGTGTCCGCTTTCTCTTCCTTGGGAAAGCTGCG 2100  
Db 1622 TGTTCGACCCCTGCTACCGGATACCGTGTCCGCTTTCTCTTCCTTGGGAAAGCTGCG 1563  
QY 2101 GCTTTCTCATAGCTCAAGCTGTAGGTATCTCAGTTGAGGTGATGCTGCTCAAGCT 2160  
Db 1562 GCTTTCTCATAGCTCAAGCTGTAGGTATCTCAGTTGAGGTGATGCTGCTCAAGCT 1503  
QY 2161 GGGCTGTGACAGAACCCCGCTTACCGCCGAGCCGCTGCGCTTATCCGGTAACTATCG 2220  
Db 1502 GGGCTGTGACAGAACCCCGCTTACCGCCGAGCCGCTGCGCTTATCCGGTAACTATCG 1443  
QY 2221 TCTTGAATTCACCCCGTGAAGACAGACTTATCCGCACTGAGACAGCTAGTAAAG 2280  
Db 1442 TCTTGAATTCACCCCGTGAAGACAGACTTATCCGCACTGAGACAGCTAGTAAAG 1383  
QY 2281 GATTAGCAGACGAGGTATGAGGCGGTGCTACAGAGTTTGAAGGTGAGCTTAACCTA 2340  
Db 1382 GATTAGCAGACGAGGTATGAGGCGGTGCTACAGAGTTTGAAGGTGAGCTTAACCTA 1323  
QY 2341 CGGCTACCTGAAGAGACAGTATTTGTATCTGCGCTGCTGTAAGCCATTAACCTCGG 2400  
Db 1322 CGGCTACCTGAAGAGACAGTATTTGTATCTGCGCTGCTGTAAGCCATTAACCTCGG 1263  
QY 2401 AAAAAGGTTGTAGCTCTTATCCGCAAAACCAACCGCTGTAGCGGTGTTTTT 2460  
Db 1262 AAAAAGGTTGTAGCTCTTATCCGCAAAACCAACCGCTGTAGCGGTGTTTTT 1203  
QY 2461 TGTTTGACAGACAGATTAACGCGCAAAAGAGATCTCAAGAGATCTTGAATCTT 2520  
Db 1202 TGTTCGAAACAGCAATTAACGCGCAAAAGAGATCTCAAGAGATCTTGAATCTT 1143  
QY 2521 TTCTACGGGGTGTGACGCTCAGTGAACGAAACCTCACTTAAGGATTTTGTATGAG 2580  
Db 1142 TTCTACGGGGTGTGACGCTCAGTGAACGAAACCTCACTTAAGGATTTTGTATGAG 1083  
QY 2581 ATTATCAAAAGAGATCTTCACTAGATCTTTTAAATTAAGATTTTAAATCAAT 2640  
Db 1082 ATTATCAAAAGAGATCTTCACTAGATCTTTTAAATTAAGATTTTAAATCAAT 1023  
QY 2641 CTAAAGTATATATGATTAACCTTGTCTGACATTAACCAATGCTTAAATCAAGTACGAC 2700  
Db 1022 CTAAAGTATATATGATTAACCTTGTCTGACATTAACCAATGCTTAAATCAAGTACGAC 963  
QY 2701 TATCTCAGGATCTGTCTAATTTGCTTATCAATCAATGCTTAAATCAAGTACGAC 2760  
Db 962 TATCTCAGGATCTGTCTAATTTGCTTATCAATCAATGCTTAAATCAAGTACGAC 903  
QY 2761 AACTACGATACGGAGGGCTTAAACATTTGGCCCAAGTCTGCAATGAATCCGAGAACCC 2820  
Db 902 AACTACGATACGGAGGGCTTAAACATTTGGCCCAAGTCTGCAATGAATCCGAGAACCC 843  
QY 2821 ACGCTCACCGGCTTCAGATTTACAGCAATAAACCAAGCCGAGAGGCGGAGGCGAG 2880  
Db 842 ACGCTCACCGGCTTCAGATTTACAGCAATAAACCAAGCCGAGAGGCGGAGGCGAG 783  
QY 2881 AAGTGTCTCTGCAACTTATCCGCTTCAATCAAGTCTAATTAATTTGCGGGAAGCTAG 2940



Db 782 AAGTGTCTGTGCACTTATCCGCTCCATCCAGTCTATTAATTGTCGGGAAGCTAG 723  
Qy 241 AGTAAGTATGTTCCAGCTTAATAGTTGGCGAAGCTTGTGCCATGTCTAAGAGCATGT 3010  
Db 722 AGTAAGTATGTTCCAGCTTAATAGTTGGCGAAGCTTGTGCCATGTCTAAGAGCATGT 663  
Qy 3001 GGTGTCACGCTGCTGTGTAAGGCTTCAATTCAGCTCCGGTCCCAACATCAAGGCG 3060  
Db 662 GGTGTCACGCTGCTGTGTAAGGCTTCAATTCAGCTCCGGTCCCAACATCAAGGCG 603  
Qy 3061 AGTACATGATCCGCCATGTTGTGCAAAAAGCGGTAGCTCTTCCGTCCTCCGATGCT 3120  
Db 602 AGTACATGATCCGCCATGTTGTGCAAAAAGCGGTAGCTCTTCCGTCCTCCGATGCT 543  
Qy 3121 TGTCAAGATGATGTTGGCCGAGTGTATCACTCATGTTATGCGACATGCTAATTC 3180  
Db 542 TGTCAAGATGATGTTGGCCGAGTGTATCACTCATGTTATGCGACATGCTAATTC 483  
Qy 3181 TCTTACTGTCAATGCGATCCGTAAGATGCTTTCTGTGCTGTGATCTCAACCAAGTC 3240  
Db 482 TCTTACTGTCAATGCGATCCGTAAGATGCTTTCTGTGCTGTGATCTCAACCAAGTC 423  
Qy 3241 ATTCTGAATATGATGCGCGCACCGAGTGTCTTGGCCGCTCAATACGAGATAA 3300  
Db 422 ATTCTGAATATGATGCGCGCACCGAGTGTCTTGGCCGCTCAATACGAGATAA 363  
Qy 3301 TACCGCGCCACATAGCAGAACCTTTAAAGTCTCATTTGAAAACGTTCTTGGGGCG 3360  
Db 362 TACCGCGCCACATAGCAGAACCTTTAAAGTCTCATTTGAAAACGTTCTTGGGGCG 303  
Qy 3361 AAAAATCTGCAAGATCTTAAAGCGCTGTGAGATCCAGTTCGATGTAACCACTCCGTCACC 3420  
Db 302 AAAAATCTGCAAGATCTTAAAGCGCTGTGAGATCCAGTTCGATGTAACCACTCCGTCACC 243  
Qy 3421 CAATGATCTTCAAGATCTTAAAGCGCTGTGAGATCCAGTTCGATGTAACCACTCCGTCACC 3480  
Db 242 CAATGATCTTCAAGATCTTAAAGCGCTGTGAGATCCAGTTCGATGTAACCACTCCGTCACC 183  
Qy 3481 GCAAAATGCGCAAAAAGGAAATGAGGCGACACGGAATGTTGAATATCTCATATCTTT 3540  
Db 182 GCAAAATGCGCAAAAAGGAAATGAGGCGACACGGAATGTTGAATATCTCATATCTTT 123  
Qy 3541 CCTTTTCAATATATGAAACATTTATCAAGGTTATGTTCTCATAGCGGATCATATTT 3600  
Db 122 CCTTTTCAATATATGAAACATTTATCAAGGTTATGTTCTCATAGCGGATCATATTT 63  
Qy 3601 TGAATGATATTTAGAAAAATAAACAATAGGGGTTCCGCGCATATTTCCCGAAAAAGTSCC 3660  
Db 62 TGAATGATATTTAGAAAAATAAACAATAGGGGTTCCGCGCATATTTCCCGAAAAAGTSCC 3  
Qy 3661 AC 3662  
Db 2 AC 1

RESULT 10  
AX306330 4001 bp DNA linear PAT 11-DEC-2001  
LOCUS AX306330  
DEFINITION Sequence 13 from Patent WO018121.  
ACCESSION AX306330  
VERSION AX306330.1 GI:17645569  
KEYWORDS  
SOURCE  
ORGANISM  
synthetic construct  
synthetic construct  
artificial sequences.  
REFERENCE  
1  
AUTHORS Plaetnick, G., Renard, J.P. and Bogaert, T.  
TITLE Vector constructs  
JOURNAL Patent: WO 018121-A 13 22-NOV-2001;  
Devgen NV (BE)  
LOCATION/Qualifiers  
FEATURES  
Source 1..4001  
/organism="synthetic construct"

ORIGIN  
Query Match 71.3%; Score 2609.4; DB 6; Length 4001;  
Best Local Similarity 83.1%; Pred. No. 0;  
Matches 3069; Conservative 0; Mismatches 591; Indels 35; Gaps 7;  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="Plasmid pGN59A"

Qy 1 CTAATTTGTAAGCGTTAATATTTTGTAAATTTGCGCTAAATTTTGTAAATCAAGTC 60  
Db 63 CGAATTTGTAAGCGTTAATATTTTGTAAATTTGCGCTAAATTTTGTAAATCAAGTC 122  
Qy 61 ATTTTAAACCAATAGCGCCGAAATCCGCAAAATCCCTTAATATCAAAAGATGACCGA 120  
Db 123 ATTTTAAACCAATAGCGCCGAAATCCGCAAAATCCCTTAATATCAAAAGATGACCGA 182  
Qy 121 GATAGGGTGAAGTGTGTTCCAGTTTGGAAACAAGTCCACTATTAAGAAAGTGAAGTC 180  
Db 183 GATAGGGTGAAGTGTGTTCCAGTTTGGAAACAAGTCCACTATTAAGAAAGTGAAGTC 242  
Qy 181 CAACGTCAAAAGGCGGAAACCGTCTATCAAGGCGATGCGCCACTAGCTGAACCATGACC 240  
Db 243 CAACGTCAAAAGGCGGAAACCGTCTATCAAGGCGATGCGCCACTAGCTGAACCATGACC 302  
Qy 241 CTAATCAAGTTTGTGGGGTCAAGTCCGTTAAAGCACTAATATGAAACCTTAAGGAG 300  
Db 303 CAATCAAGTTTGTGGGGTCAAGTCCGTTAAAGCACTAATATGAAACCTTAAGGAG 362  
Qy 301 CCCCCGATTTAAGCTTGAACGGGGGAAGCGGGGGAACGTTGGCGAAGAAAGGGAAGAA 360  
Db 363 CCCCCGATTTAAGCTTGAACGGGGGAAGCGGGGGAACGTTGGCGAAGAAAGGGAAGAA 422  
Qy 361 AGCGAAAGAGCGGCGCTAGAGCGCTGAGCAAGTGAAGCGGTCAAGCTGCGCTAACAC 420  
Db 423 AGCGAAAGAGCGGCGCTAGAGCGCTGAGCAAGTGAAGCGGTCAAGCTGCGCTAACAC 482  
Qy 421 CACACCGCGCGCTTAATGCGCGCTAAGAGGCGCTCCCATTCGCCATTACGGCTGCG 480  
Db 483 CACACCGCGCGCTTAATGCGCGCTAAGAGGCGCTCCCATTCGCCATTACGGCTGCG 541  
Qy 481 CAATGTTGGGAAGGGGATCGGCGGGGCTGTCCTTAAGCGAAGCTGCGGGAAGG 540  
Db 542 CAATGTTGGGAAGGGGATCGGCGGGGCTGTCCTTAAGCGAAGCTGCGGGAAGG 601  
Qy 541 GGGATGCTGCAAGCGATTAAGTTGGTAAAGCCAGGGTTCCTCAAGTCAAGCGCTTG 600  
Db 602 GGGATGCTGCAAGCGATTAAGTTGGTAAAGCCAGGGTTCCTCAAGTCAAGCGCTTG 661  
Qy 601 TAAACGACGCGCAGTAGCGCGCT-----CGTTCAATTCAGTTTGA 646  
Db 662 TAAACGACGCGCAGTAGTAGTGAATGCACTCTAAGGGCGAATTCAGAGCTCGGTAC 721  
Qy 647 CCGGTGAAGAGCGGGCAGACTCGCGGTCGCAAAATGTGTTTACAGCGTGAAGGACAGTG 706  
Db 722 CCGGTGAAGAGCGGGCAGACTCGCGGTCGCAAAATGTGTTTACAGCGTGAAGGACAGTG 781  
Qy 707 AAGATGCTGCAAGCGTGCAGAAACAGCAGCTAGAT-----TAACCTAAGAAAGATAA 759  
Db 782 CCGGCTTACATAAGACAGATTAACAGATGCGATTTTCCGCGCTGATTTTCCGCTATTA 841  
Qy 760 TCATATTTGTAAGCTTAAAGATTAATCAATGCGTAATATG-----ACGATGGAGT 812  
Db 842 GAATATATATGATATATATACCGAAGTATGTCATAAAGAGGTGTCATGAAGCAGCG 901  
Qy 813 CTGTAATACGATCATAATAGGGGGAATGCGGATCCCGGCTCGAGGTGAGAGGT 872  
Db 902 TATTAACGATACGATTAACGACAGCACTATCAATGCTCAAGGCAATATGATGTAAT 961  
Qy 873 ATCGATAGCTTGAATATGCAATTCCTGACCGCGGGGATCACTAGTTTGAAGCGGCC 932  
Db 962 ATCTCGGCTGTAAGACACACAGATGCAAGATGAAGAGCCGCTGCTGCGTGAAGCGC 1021

QY 933 GCCACCGGATGAGCTCCAGCTTTGTTCCCTTATGAGGGTTAATTAGATCCAGTC 992  
| | | | |  
Db 1022 TGGAAAGCCGAAATATAGGAAGGATGCTGAGGTCGCCGGTTATTATGAATGAAGCC 1081  
| | | | |  
QY 993 GTCAATTTTACGAGACTA--TCTTTAGGGTTAATCTAGCTGCATCAGATCATATC 1049  
| | | | |  
Db 1082 TCTTTTGTGACGAGAAACAGGACTGTGAAATGACGTTTAAAGTTTAACTATAAAG 1141  
| | | | |  
QY 1050 GTGGGATCTTTTTCGGGCTCAATCATGCCCCAAAGTGGGCTATCGGCGATCGGAGAG 1109  
| | | | |  
Db 1142 AGAGAGCCGTTATCTGTGTGTGAGATGACAGAGATATTAATGACAGGCCGGGGG 1201  
| | | | |  
QY 1110 GAGAAGCCCGTCCCTTTTCCCGGAGTTGAAGCGGCATGGAAGAAGTTTCCGAGAT 1169  
| | | | |  
Db 1202 ACGATGTGATCCCGCTGGCC-AGTGACGTCCTTAAAGCAATTAAGTCCCGTGAAAC 1260  
| | | | |  
QY 1170 GACTGTGCTG-CAATTGAGCTTGAAGCAAAACGACGTTTACATGATGATTCGGGAAG 1227  
| | | | |  
Db 1261 TTTTACCGGATGTGTCATATCGGGATGAAAGCTGGGCAATGATGACACCGATATGGCA 1320  
| | | | |  
QY 1228 GTTGGCCATGACGCGCTTTTAAAGGTTAACTGTTGCTGAGGCACTGGGATCCAGTT 1287  
| | | | |  
Db 1321 GTTGGCCGTCCTCGTTATCGGGGAAAGATGCTATCTACGCCACCGGAAATGACA 1380  
| | | | |  
QY 1288 CGTGGCGCTTTCCGACACAGTTCCGATGTGACGCCGAAGCGCATCAGCAACCGCA 1347  
| | | | |  
Db 1381 TCAGAAAACGCATTTAACTGATGTTCGGGAAATTAATGACAGCTCCCTTATACACA 1440  
| | | | |  
QY 1348 ACAATACCGGCGACAGCCGGAACCTGCGCGGATGACATTAATGACAGCGGTGGG 1407  
| | | | |  
Db 1441 GCGTTTTCAGACAAATGATCTGAGGGAATCTTCCATACCTACAGCTTCGCGCTGAG 1500  
| | | | |  
QY 1408 CGTGGGATATTAATCTGACAGGAGCGGATCTGGCTGATGCGCAGAAATGACAT 1467  
| | | | |  
Db 1501 GTGCGGCGCGCACTCTTGAAGTGAAAGCTTTCGCGCTATGATGATGCTATTAACAG 1560  
| | | | |  
QY 1468 GGATACCCCGATGTTACCCGCGGGCGCGGCTGGCGTAATCAATGCTATAGCTTTCC 1527  
| | | | |  
Db 1561 CTGAGTATTTATAGTGCACCTTAATAGCTTGGCGTAATCAATGCTATAGCTTTCC 1620  
| | | | |  
QY 1528 TGTGGAATTTGTTATCCGCTCACAATTCACACAAATACGAGCCGGAAGCAATTAAGTC 1587  
| | | | |  
Db 1621 TGTGGAATTTGTTATCCGCTCACAATTCACACAAATACGAGCCGGAAGCAATTAAGTC 1680  
| | | | |  
QY 1588 TAAAGCTGGGGTGTCTATAGTGAAGCTAATCACTAATTAATGCTTGGCTACTGCC 1647  
| | | | |  
Db 1681 TAAAGCTGGGGTGTCTAATGAGCTAATCACTAATTAATGCTTGGCTACTGCC 1740  
| | | | |  
QY 1648 CGCTTTCAGTCCGGAACCTGTGTCAGCTGCACTTAATGATGAGCAACGCGCGAG 1707  
| | | | |  
Db 1741 CGCTTTCAGTCCGGAACCTGTGTCAGCTGCACTTAATGATGAGCAACGCGCGAG 1800  
| | | | |  
QY 1708 GAGAGCGGTTTGGCTATTTGGGCGCTTTCGCGCTTCGCTCACTGATCGCTGGCTC 1767  
| | | | |  
Db 1801 GAGAGCGGTTTGGCTATTTGGGCGCTTTCGCGCTTCGCTCACTGATCGCTGGCTC 1860  
| | | | |  
QY 1768 GGTGTTTGGGCTGGGCGAGCGGTTATCAGCTCACTGAAAGCGGTAATACGTTATTCAC 1827  
| | | | |  
Db 1861 GGTGTTTGGGCTGGGCGAGCGGTTATCAGCTCACTGAAAGCGGTAATACGTTATTCAC 1920  
| | | | |  
QY 1828 AGATACAGGGATTAACGAGAAAGAACTGTGAGAAAGGCGCAAAAGGCGAGAA 1887  
| | | | |  
Db 1921 AGATACAGGGATTAACGAGAAAGAACTGTGAGAAAGGCGCAAAAGGCGAGAA 1980  
| | | | |  
QY 1888 CGGTAAAGGCGCGCTTGTGCTGGCGTTTTCATAGGCTCGCGCCCTGACGAGATCA 1947  
| | | | |  
Db 1981 CGGTAAAGGCGCGCTTGTGCTGGCGTTTTCATAGGCTCGCGCCCTGACGAGATCA 2040  
| | | | |  
QY 1948 CAAAAATGACGCTCAAGTCAAGGTTGGGAAACCCGACAGACTATTAAGATACAGGC 2007  
| | | | |  
Db 2041 CAAAAATGACGCTCAAGTCAAGGTTGGGAAACCCGACAGACTATTAAGATACAGGC 2100  
| | | | |  
QY 2008 GTTTCCTGGAAGCTCCCTGTCGCTCTCTGTTCCAGCTGCGGCTTACCGGATA 2067  
| | | | |

Db 2101 GTTTCCTGGAAGCTCCCTGTCGCTCTCTGTTCCAGCTCCCTGATCCGGATA 2160  
| | | | |  
QY 2068 CCGTCCGCGCTTTCCTTCGGAAGCGTGGCGCTTTCATAGCTCAGCTGATGATA 2127  
| | | | |  
Db 2161 CCGTCCGCGCTTTCCTTCGGAAGCGTGGCGCTTTCATAGCTCAGCTGATGATA 2220  
| | | | |  
QY 2128 TCTCAGTTGGGTGATGCTGTTGCTCAAGAGTGGGCTGTGCAAGAACCCCGTTCA 2187  
| | | | |  
Db 2221 TCTCAGTTGGGTGATGCTGTTGCTCAAGAGTGGGCTGTGCAAGAACCCCGTTCA 2280  
| | | | |  
QY 2188 GCCGACCGGTGGCGCTTATCCGTTAATATATGTCCTTGAATCAACCGGTAAGACAGA 2247  
| | | | |  
Db 2281 GCCGACCGGTGGCGCTTATCCGTTAATATATGTCCTTGAATCAACCGGTAAGACAGA 2340  
| | | | |  
QY 2248 CTTATCGCACTGCGACAGCACTGATGATGAGATTAAGAGAGATGATGAGCGG 2307  
| | | | |  
Db 2341 CTTATCGCACTGCGACAGCACTGATGATGAGATTAAGAGAGATGATGAGCGG 2400  
| | | | |  
QY 2308 TGTATGAGGTTCTTGAATGCTGCTTAACTACCGGCTACACTAGAAAGACATATTTGG 2367  
| | | | |  
Db 2401 TGTATGAGGTTCTTGAATGCTGCTTAACTACCGGCTACACTAGAAAGACATATTTGG 2460  
| | | | |  
QY 2368 TATCTGCGCTGTGTAAGCCAGTTACTTGGAAGAAAGATGATGCTTGAATCGG 2427  
| | | | |  
Db 2461 TATCTGCGCTGTGTAAGCCAGTTACTTGGAAGAAAGATGATGCTTGAATCGG 2520  
| | | | |  
QY 2428 CAAACAAACCCGCTGTGATGCGGTGTTTTTTTGTGCAAGACAGATTAACGCGAG 2487  
| | | | |  
Db 2521 CAAACAAACCCGCTGTGATGCGGTGTTTTTTTGTGCAAGACAGATTAACGCGAG 2580  
| | | | |  
QY 2488 AAAAAAGATCTCAAGAAATCTTGTATCTTTTCAACGGGCTGACGCTCATGTGAA 2547  
| | | | |  
Db 2581 AAAAAAGATCTCAAGAAATCTTGTATCTTTTCAACGGGCTGACGCTCATGTGAA 2640  
| | | | |  
QY 2548 GGAAGACCTGTAAGGATTTTGTGATGATTAAGAAAGATCTTACCTAGAT 2607  
| | | | |  
Db 2641 GGAAGACCTGTAAGGATTTTGTGATGATTAAGAAAGATCTTACCTAGAT 2700  
| | | | |  
QY 2608 CTTTTTAAATTAATGAATTTTAAATCAATCTAATATATATGATTAATTTGGTC 2667  
| | | | |  
Db 2701 CTTTTTAAATTAATGAATTTTAAATCAATCTAATATATATGATTAATTTGGTC 2760  
| | | | |  
QY 2668 TGAAGTTACCATGCTTAATCAGTGAAGCACTATCTCAGGATCTGCTATTTGCTTC 2727  
| | | | |  
Db 2761 TGAAGTTACCATGCTTAATCAGTGAAGCACTATCTCAGGATCTGCTATTTGCTTC 2820  
| | | | |  
QY 2728 ATCATAGTTGCTGATCTCCCGTGTGATGATTAATCAATGATGAGGCTTACATC 2787  
| | | | |  
Db 2821 ATCATAGTTGCTGATCTCCCGTGTGATGATTAATCAATGATGAGGCTTACATC 2880  
| | | | |  
QY 2788 TGGCCCCAGTGTGCAATATACCGCGAGACCAAGCTCAGCGCTCAGATTTATCAGC 2847  
| | | | |  
Db 2881 TGGCCCCAGTGTGCAATATACCGCGAGACCAAGCTCAGCGCTCAGATTTATCAGC 2940  
| | | | |  
QY 2848 AATTAACAGCCAGCGGAGGAGCGAGAGGAGTGTCTGCAATTTATCCGCTC 2907  
| | | | |  
Db 2941 AATTAACAGCCAGCGGAGGAGCGAGAGGAGTGTCTGCAATTTATCCGCTC 3000  
| | | | |  
QY 2908 CATCAGTCTATTAATTTTTCGCGGAAGCTAGATGATGATTCGCAATTAATGTTT 2967  
| | | | |  
Db 3001 CATCAGTCTATTAATTTTTCGCGGAAGCTAGATGATGATTCGCAATTAATGTTT 3060  
| | | | |  
QY 2968 GCGCAAGCTTGTGCAATTTGCTACAGCATCTGATGATCAGCTGCTGTTTGGTATGCG 3027  
| | | | |  
Db 3061 GCGCAAGCTTGTGCAATTTGCTACAGCATCTGATGATCAGCTGCTGTTTGGTATGCG 3120  
| | | | |  
QY 3028 TTTCAATCAGCTCGGTTCCCAAGATCAAGCGAGTTACATATCCCATGTTGGCA 3087  
| | | | |  
Db 3121 TTTCAATCAGCTCGGTTCCCAAGATCAAGCGAGTTACATATCCCATGTTGGCA 3180  
| | | | |  
QY 3088 AAAAGCGTTAGCTCCCTGGTCTCCGATCGTTGCAAGTAATGCGCAGTGT 3147  
| | | | |

Db 3181 AAAAGCGTTAGTCTCTCGGTCTCCGATGTTGTGAGAAAGTAAGTTGGCCGAGTGT 3240  
Qy 3148 ATCACTAAGGTTATGGAGACATGCAATATTTCTTACTGTGATGCGATCCGTAAGTG 3207  
Db 3241 ATCACTAAGGTTATGGAGACATGCAATATTTCTTACTGTGATGCGATCCGTAAGTG 3300  
Qy 3208 CTTTCTGTGACTGAGTACTCAACCAAGTCATTTGAGAAATGATGATGCGGCGAC 3267  
Db 3301 CTTTCTGTGACTGAGTACTCAACCAAGTCATTTGAGAAATGATGATGCGGCGAC 3360  
Qy 3368 GAGTTGCTCTTCCCGCGGTCATACCGGATTAATACCGCGCATAGCAGAACTTTAA 3327  
Db 3361 GAGTTGCTCTTCCCGCGGTCATACCGGATTAATAGTATAGACATAGCAGAACTTTAA 3420  
Qy 3328 AGTGTCTATCATTTGGAAAAAGTTCTTGGGGCGAAAAATCTCAGAGATCTTACGCTGT 3387  
Db 3421 AGTGTCTATCATTTGGAAAAAGTTCTTGGGGCGAAAAATCTCAGAGATCTTACGCTGT 3480  
Qy 3388 GAGATCCAGTTCAGATGTAACCCACTCGTGCACCCCAAGTATCTTCAGCATCTTTTACTT 3447  
Db 3481 GAGATCCAGTTCAGATGTAACCCACTCGTGCACCCCAAGTATCTTCAGCATCTTTTACTT 3540  
Qy 3448 CACCAAGCGTTTCTGGGTGAGCAAAAAACAGAAAGCAAAATGCCGAAAAAAGGAATAG 3507  
Db 3541 CACCAAGCGTTTCTGGGTGAGCAAAAAACAGAAAGCAAAATGCCGAAAAAAGGAATAG 3600  
Qy 3508 GGGCAGCAGGAAATGTTGAATATCTATCTCTTCTTTTCAATTTATTTGAGCATTTA 3567  
Db 3501 GGGCAGCAGGAAATGTTGAATATCTATCTCTTCTTTTCAATTTATTTGAGCATTTA 3660  
Qy 3568 TCAGGGTATTTGTCTCATGAGCGGATACATATTTGAATGATTTGAAGAAAAATTAACAAT 3627  
Db 3661 TCAGGGTATTTGTCTCATGAGCGGATACATATTTGAATGATTTGAAGAAAAATTAACAAT 3720  
Qy 3628 AGGGGTTCCGCGCACATTTCCCGAAAAAGTCCAC 3662  
Db 3721 AGGGGTTCCGCGCACATTTCCCGAAAAAGTCCAC 3755

RESULT 11  
AF179627 5540 bp DNA linear SYN 22-NOV-1999  
LOCUS AF179627  
DEFINITION Expression vector pCENEX645, complete sequence.  
ACCESSION AF179627  
VERSION AF179627.1 GI:6457303  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
FEATURES  
SOURCE

1 (bases 1 to 5540)  
Bugster,H.P., Bartesch,S., Wurgler,F.B. and Sengstag,C.  
Functional co-expression of human oxidoreductase and cytochrome  
P450 1A1 in Saccharomyces cerevisiae results in increased EROD  
activity  
Biochem. Biophys. Res. Commun. 185 (2), 641-647 (1992)  
1610357  
2 (bases 1 to 5540)  
Sengstag,C. and Paladino,G.  
A collection of transformed Saccharomyces cerevisiae strains as an  
unlimited source of human enzymes  
Biomol. 97 (6), 2-7 (1997)  
3 (bases 1 to 5540)  
Sengstag,C.  
Direct Submission  
Submitted (24-AUG-1999) Genetics, Institute of Toxicology, ETH  
Zurich, Schorenstrasse 16, Schwerzenbach 8603, Switzerland  
Location/Qualifiers  
1..5540  
/organism="Expression vector pCENEX645"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:106613"  
/note="yeast expression vector pCENEX645 (alias pCS645)"

containing truncated GAPDH promoter for constitutive  
expression of cDNAs in yeast"  
417..1220  
/gene="URA3"  
417..1220  
/gene="URA3"  
/note="yeast selection marker"  
/codon\_start=1  
/product="URA3"  
/protein\_id="AA09466.1"  
/db\_xref="GI:6457304"  
/translation="MSKATYKERAPATPSPVAAKLFINIHEKOTNLCAIDVRYTEL  
LEIYEAAGPILCLTKHVDILTDESMGTVPKALASAKXNFIIFEDRKAFADICNTYK  
LOYSAGVRIAPADITTAAGVYVSGVIGLAKQAAEVTKEPRLILALISCGSIS  
TGEYTKGTVDIASKDQFVIGFTIQRDMGDESDYDLINTPGVGLDDKCTQQR  
TVDDVSTGSDIILVGRGLPAKGDADAEERYRKAGWEAVLRGQGN"  
2310..2508  
/note="truncated GAPDH promoter"  
complement (3958..4818)  
/gene="Amp"  
complement (3958..4818)  
/gene="Amp"  
/note="Amp"  
marker"  
/codon\_start=1  
/product="Amp"  
/protein\_id="AA09467.1"  
/db\_xref="GI:6457305"  
/translation="MSIOHFRVALIPEFAFCLVPAHPDLVFKDCEDQLGARVGY  
IELDNSGKILIESFRPEPRFPMSTFVILGAVLFPIIDGOEOLGRIHYSNDLYE  
YSPVTEKHLIDMGTVRELGSAITMSDNTANILITIGPKELTAFLHMGDVTRL  
DRPEPELNEAIPNDERDTTPVMAATTLRLKLTGSELLTASROOLIMMEDKRYAGPL  
LRSLPMPWFTIADSKSAGERSKGIINALGPDGKPSRIVVIYTTGSOATIDERNRQIA  
EIGASLIKHW"

ORIGIN  
Query Match 69.9%; Score 2558.8; DB 12; Length 5540;  
Best Local Similarity 82.8%; Pred. No. 0;  
Matches 3035; Conservative 0; Mismatches 557; Indels 72; Gaps 7;  
Qy 3 AAATTGTAAGCGTTAATATTTTGTAAAAATTCGCGTTAAATTTTGTAAATCAGTCAT 62  
Db 1351 AAATTGTAAGCGTTAATATTTTGTAAAAATTCGCGTTAAATTTTGTAAATCAGTCAT 1410  
Qy 63 TTTTAAACAATAGCGCGAAATTCGCAAAATCCCTTATTAATCAAAGAATAGACCGAGA 122  
Db 1411 TTTTAAACAATAGCGCGAAATTCGCAAAATCCCTTATTAATCAAAGAATAGACCGAGA 1470  
Qy 123 TAGGGTGAAGTGTTCGAGTTTGGAAACAAGTCCACTATTAAAGAGTGGACTCCA 182  
Db 1471 TAGGGTGAAGTGTTCGAGTTTGGAAACAAGTCCACTATTAAAGAGTGGACTCCA 1530  
Qy 183 ACGTCAAGGCGCAAAAACGCTCATAGGGCGATGCGCACTACGTGAACCATCACCT 242  
Db 1531 ACGTCAAGGCGCAAAAACGCTCATAGGGCGATGCGCACTACGTGAACCATCACCT 1590  
Qy 243 AATCAAGTTTTCGAGGTGCAAGTCCGTAAGACATTAATCGAACCTTAAGGAGCC 302  
Db 1591 AATCAAGTTTTCGAGGTGCAAGTCCGTAAGACATTAATCGAACCTTAAGGAGCC 1650  
Qy 303 CCGCATTTAAGCTTGAACGGGGAAGCGCGCAACGTGGGGAAGAAGGAGGAAG 362  
Db 1651 CCGCATTTAAGCTTGAACGGGGAAGCGCGCAACGTGGGGAAGAAGGAGGAAG 1710  
Qy 363 CGAAGAGCGGCGCTAGAGCGCTGCAAGTGAAGCGTCAACGCTGCGCTAACCA 422  
Db 1711 CGAAGAGCGGCGCTAGAGCGCTGCAAGTGAAGCGTCAACGCTGCGCTAACCA 1770  
Qy 423 CACCGCGCGCTTAATGCGCGCTACAGGCGCGTCCATTCGCGCATTCAGGCTGC 479  
Db 1771 CACCGCGCGCTTAATGCGCGCTACAGGCGCGTCCATTCGCGCATTCAGGCTGC 1830  
Qy 480 GCAACTGTTGGGAAGGCGAATCGGTGCGGCGCTTTCGCTATTACGCAAGTGGCGAAG 539

Db 1831 GCACTGTTGGGAAAGGCGATCGGTGGGCGCTTCGCTATTAGCCAGCTGGGAAAG 1890  
Qy 540 GGGATGTGCTGCAAGCGATTAAAGTGGGTAAAGCAGAGTTTTCCTCCAGTCACAGCGTT 599  
Db 1891 GGGGATGTGCTGCAAGCGATTAAAGTGGGTAAAGCAGAGTTTTCCTCCAGTCACAGCGTT 1950  
Qy 600 GTAAAGCAGCGCGCATGAGCGCGCTCGTTCAATTCACGTTTGTAAACCCCGTAGAGACG 659  
Db 1951 GTAAAGCAGCGCATGAG-----ATTGTAAATACGACTCACTAATAGGCCAATTTGG 2001  
Qy 660 GCGAGACTCGCGGTGCAAAATGTGTTTACAGGCTATAGAGCAATGAAGTCTGCACA 719  
Db 2002 AGCTCCACCGCGGTG-----GCGGCGCTCTAGTCACGCTCTCC 2042  
Qy 720 CGCTGCAAGAACGCGAGTAACTTAACTTAAAGATTAATCATATTTGTGACGTACGTTA 779  
Db 2043 TTATGGACTCTGCACTTAAGAAAGCAGCCAGTAGTGTGAAGCGCTTGACACCGCC 2102  
Qy 780 AAGATTAATCAGCTGTAATTAATGAAGCATGGGATCTGTAAATACACTATAGGCGAA 839  
Db 2103 GCGCAAGAAATGTGTGATGCAAGAGATGCGGCCCAAGTCCCGCGCACGCGGCGCT 2162  
Qy 840 TTGGTACCGGAGCCCCCTCGAGTGCAGCGTATGCAATGCTGTAAATCAATTCGCG 899  
Db 2163 GCAACATACCAAGCGCGAAACAGAGCGCTCATAGCGCCGAAGTGGAGCCGATCTTC 2222  
Qy 900 CAGCCCGGGGATCACTAGTTCTAGAGCGCGCCGACGCGGTGAGCTCCAGCTTTTG 959  
Db 2223 CATGCGTGTATGTGCGGATATAGCGCCGACACGCTGTGCGCGGTGTATGCGG 2282  
Qy 960 TTCCCTTAAAGAGGTTAATTAGATCCATGCTCAATTTTACAGACATATCTTTCTA 1019  
Db 2283 GCAACATGCGTCCGGGTAGAGATCTCTGAAAAAAGTTGAAACGATTCCTG 2342  
Qy 1020 GGGTAACTAGCTGATAGATGATCATATCGTGGGTCTTTTCCGCGCTCACTATCGC 1079  
Db 2343 AATTTATTC-----CCTTACTGACTAATTAAGTAAATAA 2376  
Qy 1080 CCAAGCTGCGCTATCTGGGATCGGGAGAGAAAGCCGCTGCTTTTCCCGAGAGTT 1139  
Db 2377 AGACGGTAGGTATGATTTGATTTCTGA-----AATCTATTTCTTAACTTTGTAAT 2431  
Qy 1140 GAAGCGCATGAAAGAGTTTGCAGAGATGCTGCTGCAATGTACGTTGACGAA 1199  
Db 2432 CTACTTTATATGTTAGTCTTTTTTTGTAGTT-----TTAAACACCAAGAACTTAAT 2482  
Qy 1200 CGCAGTTTACCATGATGATTGGGAAGGTGCGCATGCAAGCTTTTAAAGGTGAATG 1259  
Db 2483 TTGCAATTAACACATTAATTAAGATTCATGCTTTTGCAGAGGAAACATACTATTAC 2542  
Qy 1260 TTGCTGACGCGACCTGGGATACCAATTCGTCGCGCTTTTCCGACACAGTTCCGATG 1319  
Db 2543 GCTGCTTTTGCAGCAAAATATCTGCACTTTAGCGTCTCGAGGGAACATACTATTAC 2602  
Qy 1320 GTGAGCCCGAAGCGCATCAGCAACCGAACAATACCGGCAAGCGCGAATCGCGTGC 1379  
Db 2603 AACCCAGTCTATTGAGACAATGTTTGTATTAATTAATTAATTAATTAATTAATTAAT 2662  
Qy 1380 GGTGTGAGTAAATGACAGCGGTGCGGCGTGGATATTAAGTCAAGAGAGAGCGGTAT 1439  
Db 2663 AATACCAATTTTATTTTAAATTTTGCAGAAATTAATTAATTAATTAATTAATTAATTAAT 2722  
Qy 1440 CTTGCGTGAATGCGCAGAAATGGAATAGATACCCGCTGATTAACCGGCGCGG 1498  
Db 2723 CGAGGGGGGCGCGCGCGCGCGCGCTTTTGTCCCTTATGATAGGGTAAATTCGAGC 2782  
Qy 1499 TTGGGCTATCATGTGATAGCTGTTTCTGTGTGAATTTGTATCGGCTCAAAATTTCA 1558  
Db 2783 TTGGGCTATCATGTGATAGCTGTTTCTGTGTGAATTTGTATCGGCTCAAAATTTCA 2842  
Qy 1559 CACAATATGAGCGGAGCATTAAGTTAAAGCTTGGGCTGCTAATAGTAGACTAA 1618  
|||||

Db 2843 CACAATATGAGCGGAGAGCATTAAGTTAAAGCTTGGGCTGCTAATAGTAGACTAA 2902  
Qy 1619 CTCACATTAATTCGTTGCTGCTCACTGCGGCTTTCCAGTCGGAAACCTGTGTCCAG 1678  
Db 2903 CTCACATTAATTCGTTGCTGCTCACTGCGGCTTTCCAGTCGGAAACCTGTGTCCAG 2962  
Qy 1679 CTCATTAATTAATTCGCTCAACCGCGGGAAGAGCGGTGCGTATTTGGGCGCTTTC 1738  
Db 2963 CTCATTAATTAATTCGCTCAACCGCGGGAAGAGCGGTGCGTATTTGGGCGCTTTC 3022  
Qy 1739 GCTTCTGCTCACTGACTGCTGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1798  
Db 3023 GCTTCTGCTCACTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3082  
Qy 1799 CACTCAAAAGGCGGTATTAAGCTTATCAAGATCAAGGATTAAGCAGAAAGAAATG 1858  
Db 3083 CACTCAAAAGGCGGTATTAAGCTTATCAAGATTAAGGAGATTAAGCAGAAAGAAATG 3142  
Qy 1859 TGAGCAAAAGCGCAGCAAAAGCGCAAGAACGTTAAAGCGCGTTCGTCGTTTTC 1918  
Db 3143 TGAGCAAAAGCGCAGCAAAAGCGCAAGAACGTTAAAGCGCGTTCGTCGTTTTC 3202  
Qy 1919 CATAGGCTCGGCGGCTTACCGGATACCTGTGCGCTTTTCCCTTCCGGAAGCGT 2038  
Db 3203 CATAGGCTCGGCGGCTTACCGGATACCTGTGCGCTTTTCCCTTCCGGAAGCGT 3262  
Qy 1979 AACCCGACAGCATTAAGATTAACAGCGCTTTCCTGGAAGCTCCTGTGCGCT 2038  
Db 3263 AACCCGACAGCATTAAGATTAACAGCGCTTTCCTGGAAGCTCCTGTGCGCT 3322  
Qy 2039 CCTGTCCGACCGCGCTTACCGGATACCTGTGCGCTTTTCCCTTCCGGAAGCGT 2098  
Db 3323 CCTGTCCGACCGCGCTTACCGGATACCTGTGCGCTTTTCCCTTCCGGAAGCGT 3382  
Qy 2099 GCGCTTTCTCATAGCTCAAGCTGATGATCTCAGTTGCGTGAAGTCTTTCCTCAAG 2158  
Db 3383 GCGCTTTCTCATAGCTCAAGCTGATGATCTCAGTTGCGTGAAGTCTTTCCTCAAG 3442  
Qy 2159 CTGGGCTGTGTGACGAAGCCCGGTTACCGGACCGCTGCGCTTATCCGTAATCTAT 2218  
Db 3443 CTGGGCTGTGTGACGAAGCCCGGTTACCGGACCGCTGCGCTTATCCGTAATCTAT 3502  
Qy 2219 CGTCTTGAATCAACCGGTAAGACAGACTTATGCGCATGGAGCGCACTGTGTAC 2278  
Db 3503 CGTCTTGAATCAACCGGTAAGACAGACTTATGCGCATGGAGCGCACTGTGTAC 3562  
Qy 2279 AGGATTAAGCAGAGCGATTAAGCGGCTGTACAGATTTCTTGAAGTGTGCGCTTAC 2338  
Db 3563 AGGATTAAGCAGAGCGATTAAGCGGCTGTACAGATTTCTTGAAGTGTGCGCTTAC 3622  
Qy 2339 TACGGCTACATGAAGACAGATTTTGTATCTGCGCTGTGAGGCGCAATTAACCTTC 2398  
Db 3623 TACGGCTACATGAAGACAGATTTTGTATCTGCGCTGTGAGGCGCAATTAACCTTC 3682  
Qy 2399 GGAAGAAAGTGTGATCTTGAATCCGCGAAACAAACAGCGCTGTGAGCGGTGTT 2458  
Db 3683 GGAAGAAAGTGTGATCTTGAATCCGCGAAACAAACAGCGCTGTGAGCGGTGTT 3742  
Qy 2459 TTTGTTTCCAGCAGCAGATTAAGCGCAGAAAGAAAGATCTCAAGAAATCCTTTGATC 2518  
Db 3743 TTTGTTTCCAGCAGCAGATTAAGCGCAGAAAGAAAGATCTCAAGAAATCCTTTGATC 3802  
Qy 2519 TTTTCTACGCGGTGTGAGCTCAAGTGAAGAAACCTCAAGTTAAGGATTTTGTCTATG 2578  
Db 3803 TTTTCTACGCGGTGTGAGCTCAAGTGAAGAAACCTCAAGTTAAGGATTTTGTCTATG 3862  
Qy 2579 AGATTAATCAAAAGATTTTCACTAGATCTTTTAAATTAAGATTTTAAATCA 2638  
Db 3863 AGATTAATCAAAAGATTTTCACTAGATCTTTTAAATTAAGATTTTAAATCA 3922  
Qy 2639 ATCTAAAGTATATATAGTAACTTGTGTGACAGTTACCAATGCTTAATCAGTAGGCA 2698  
Db 3923 ATCTAAAGTATATATAGTAACTTGTGTGACAGTTACCAATGCTTAATCAGTAGGCA 3982  
|||||

QY 2699 CCTATCTCAGGATCTCTCTATTTCCTTCATCCATAGTTCCTGCTCCCGTCTGTAG 2758  
 Db 3883 CCTATCTCAGGATCTCTCTATTTCCTTCATCCATAGTTCCTGCTCCCGTCTGTAG 4042  
 QY 2759 ATAACTACGATACGGAGGGGCTTACCATCTGAGCCCGAGTGTGCAATGATACCGGAGAC 2818  
 Db 4043 ATAACTACGATACGGAGGGGCTTACCATCTGAGCCCGAGTGTGCAATGATACCGGAGAC 4102  
 QY 2819 CCAGGCTCAGCGGCTCCAGATTTATCAGCAATAAACAGCCAGCCGAGAGGGCCGAGGC 2878  
 Db 4103 CCAGGCTCAGCGGCTCCAGATTTATCAGCAATAAACAGCCAGCCGAGAGGGCCGAGGC 4162  
 QY 2879 AGAAGTGTCTCTGCAACTTTATCCGCTCCATCCAGTCTTATTAATTTGTCGGGAACT 2938  
 Db 4163 AGAAGTGTCTCTGCAACTTTATCCGCTCCATCCAGTCTTATTAATTTGTCGGGAACT 4222  
 QY 2939 AGAGTAAGTATGTTGCCAGTTAATAGTTGGCAAGTTGTTGCCATTTGTACAGGATC 2998  
 Db 4223 AGAGTAAGTATGTTGCCAGTTAATAGTTGGCAAGTTGTTGCCATTTGTACAGGATC 4282  
 QY 2999 GTGGTGTACGCTCGTCTGTTGGTAAGCTTCATTCAGCTCCGTTCCCAAGATCAAG 3058  
 Db 4283 GTGGTGTACGCTCGTCTGTTGGTAAGCTTCATTCAGCTCCGTTCCCAAGATCAAG 4342  
 QY 3059 CGAGTTACATGATCCCCCATGTTGTGCAAAAAAGGGGTTAGTCTCTTGCTCTCCATC 3118  
 Db 4343 CGAGTTACATGATCCCCCATGTTGTGCAAAAAAGGGGTTAGTCTCTTGCTCTCCATC 4402  
 QY 3119 GTTGTGCAAGTAAGTTGGCCGAGTGTATCATCTCATGTTATGGCAGCATGATAT 3178  
 Db 4403 GTTGTGCAAGTAAGTTGGCCGAGTGTATCATCTCATGTTATGGCAGCATGATAT 4462  
 QY 3179 TCTCTTACTGTCATGCCATCCGTAGATGCTTTCTGTGACTGTGTAGTACTCAACCA 3238  
 Db 4463 TCTCTTACTGTCATGCCATCCGTAGATGCTTTCTGTGACTGTGTAGTACTCAACCA 4522  
 QY 3239 TCATTTGCAAGTAAGTTGGCCGAGTGTATCATCTCATGTTATGGCAGCATGATAT 3298  
 Db 4523 TCATTTGCAAGTAAGTTGGCCGAGTGTATCATCTCATGTTATGGCAGCATGATAT 4582  
 QY 3299 AATAACCGGCGCATAGCAGAACTTTAAGAGTCTCATGTTGAAAAAGTTCTTCCGGG 3358  
 Db 4583 AATAACCGGCGCATAGCAGAACTTTAAGAGTCTCATGTTGAAAAAGTTCTTCCGGG 4642  
 QY 3359 CGAAAACTCTCAAGATCTTACCGCTGTGAGATCCAGTGTGATGTAACCACTCGTCA 3418  
 Db 4643 CGAAAACTCTCAAGATCTTACCGCTGTGAGATCCAGTGTGATGTAACCACTCGTCA 4702  
 QY 3419 CCCAACTGATCTTCAAGATCTTCTTCAACAGCGTTTCTGGGTAGCAAAAAACAGA 3478  
 Db 4703 CCCAACTGATCTTCAAGATCTTCTTCAACAGCGTTTCTGGGTAGCAAAAAACAGA 4762  
 QY 3479 AGGCAAAATGCGGCAAAAAAGGAATAGGGCGACAGGAAATGTGAAATCTCATCTC 3538  
 Db 4763 AGGCAAAATGCGGCAAAAAAGGAATAGGGCGACAGGAAATGTGAAATCTCATCTC 4822  
 QY 3539 TTCTTTTTCATATTTATTTGAAGATTTATCAGGGTATTTGTCTCATGAGCGGATACA 3598  
 Db 4823 TTCTTTTTCATATTTATTTGAAGATTTATCAGGGTATTTGTCTCATGAGCGGATACA 4882  
 QY 3599 TTTGAATGATTTTGAATAATTAACAAATAGGGGTTCCGCGACATTTCCCGAAAAAGTG 3658  
 Db 4883 TTTGAATGATTTTGAATAATTAACAAATAGGGGTTCCGCGACATTTCCCGAAAAAGTG 4942  
 QY 3659 CCAC 3662  
 Db 4943 CCAC 4946

RESULT 12  
 ASPGREEN2/c  
 LOCUS ASPGREEN2 3633 bp DNA linear SYN 15-MAY-1998

DEFINITION Artificial sequences, plasmid vector pGreen-2.  
 ACCESSION Y09374  
 VERSION Y09374.1 GI:1684627  
 KEYWORDS biatEM gene; lacZ-Phoc fusion gene; multiple cloning site; T3 promoter; T7 promoter.  
 SOURCE ORGANISM synthetic construct  
 ORGANISM artificial sequences.  
 REFERENCE 1 Thaller,M.C., Berlucchi,F., Schippa,S., Selan,L. and Rossolini,G.M. Bacterial acid phosphatase gene fusions useful as targets for cloning-dependent insertional inactivation Biotechnol. Prog. 14 (2), 241-247 (1998)  
 MEDLINE 98215688  
 PUBMED 9548775  
 REFERENCE 2 (bases 1 to 3633)  
 AUTHORS Rossolini,G.M.  
 TITLE Direct Submission  
 JOURNAL Submitted (11-NOV-1996) G.M. Rossolini, Dip. Biologia Molecolare-Sez., Microbiologia, Univ. di Siena, via Laterina N.8, I- 53100 Siena, ITALY  
 COMMENT Related sequences: X64444 and X52328.  
 FEATURES Location/Qualifiers  
 1..3633  
 /organism="synthetic construct"  
 /mol\_type="other DNA"  
 /db\_xref="taxon:32630"  
 /plasmid="pGreen-2"  
 129..989  
 /gene="biatEM"  
 129..989  
 /gene="biatEM"  
 /codon\_start=1  
 /transl\_table=11  
 /protein\_id="CAA70549.1"  
 /db\_xref="GI:1684628"  
 /translation="MSIQHPPVALIPFAFLPVPAPBETLVKVKDAEOLGARVY IEIDLSNGKILBSFRPBRPMSSTFVLCGLVSLRIDGEOLEGRIRYSQNDLVE YSPVTEKHLDGKTVRELCSAAITMSDNTANLLTTIGPKETLAFHMGDHTVL DRWPELNEBAIPNDERDTMPVMAATLRLLTGLGELTLASRQQLIDMRADKVAQPL LRSALPGAFWFIADKSGAGERSGIIALPDGKPSRIVIVYTTGSOATDERNROIA EIGSLIKHW"  
 2147..2162  
 /note="pUC M13/Reverse Sequencing primer recognition site"  
 2154..3104  
 /gene="lacZ-Phoc"  
 2154..3104  
 /gene="lacZ-Phoc"  
 /codon\_start=1  
 /transl\_table=11  
 /protein\_id="CAA70550.1"  
 /db\_xref="GI:1684629"  
 /translation="MTWITPSAULTLTKGNKSSPVAALAEIVDPGGRNLSLS IPTSRGAPVNPASPYSSVYARSLAVLORMDENPVTGATPAGNDATTKPLVYL KNRQALDSLKLPPRPVSGISITLNDQAMTEKRMRLNTRGQAQADADLAAAGVAT .DQKGLSTNGSVPSGHTSIGWATLVLAEPVAPQDAITLERYDLOGSRIYCGHWQSD VDAARIVGSAVAVTILHSDPAFOQLAKAKOEPAKSQKDL"  
 2179..2198  
 /gene="lacZ-Phoc"  
 /note="T3 promoter"  
 2214..2316  
 /gene="lacZ-Phoc"  
 /note="multiple cloning site"  
 2325..2344  
 /gene="lacZ-Phoc"  
 /note="T7 promoter"  
 2354..2370  
 /gene="lacZ-Phoc"  
 /note="M13-20 sequencing primer recognition site"  
 2374..2390  
 /gene="lacZ-Phoc"  
 /note="M13-40 sequencing primer recognition site"



## ORIGIN

Query Match 69.7%; Score 2550.6; DB 12; Length 3633;  
Best Local Similarity 82.3%; Pred. No. 0;  
Matches 3020; Conservative 0; Mismatches 604; Indels 47; Gaps 6;

QY 1 CTAAATTTGTAAGCGTAAATATTTTGTAAATTCGCGTAAATTTTGTAAATCAAGTC 60  
DB 3633 CTAAATTTGTAAGCGTAAATATTTTGTAAATTCGCGTAAATTTTGTAAATCAAGTC 60  
QY 61 ATTTTAAACCAATAGCGCCGAATCGCAAAATCCCTATATTAATCAAAAGATAGCCGA 120  
DB 3573 ATTTTAAACCAATAGCGCCGAATCGCAAAATCCCTATATTAATCAAAAGATAGCCGA 120  
QY 121 GATAGGTTGATGTTGTTCCAGTTTGGACAAGATCCATTTAAAGAACTGACATC 180  
DB 3513 GATAGGTTGATGTTGTTCCAGTTTGGACAAGATCCATTTAAAGAACTGACATC 180  
QY 181 CAACGTCAAGGGGCAAAAACCGTCTATCAGGGCGATGGCCCACTAGCAATCAAC 240  
DB 3453 CAACGTCAAGGGGCAAAAACCGTCTATCAGGGCGATGGCCCACTAGCAATCAAC 240  
QY 241 CTAAATCAAGTTTGTGGGGTGAAGTCCGTTAAAGCACTTAATCGAAACCTTAAAGGAG 300  
DB 3393 CTAAATCAAGTTTGTGGGGTGAAGTCCGTTAAAGCACTTAATCGAAACCTTAAAGGAG 300  
QY 301 CCCCCGATTTAGAGCTTGAACGGGGAAGCCGCGAAAGTGGGGAAGAAAGGAGAA 360  
DB 3333 CCCCCGATTTAGAGCTTGAACGGGGAAGCCGCGAAAGTGGGGAAGAAAGGAGAA 360  
QY 361 AGCGAAAGAGCGGGCGCTAGGGCGCTGGCAAGTGAAGGCTACGCTGCGGTAAACAC 420  
DB 3273 AGCGAAAGAGCGGGCGCTAGGGCGCTGGCAAGTGAAGGCTACGCTGCGGTAAACAC 420  
QY 421 CACACCCGCGCGCTTAATGCGCGCTAACAGGGCGCGTCCCATTCGACATTCAGGCTGCG 480  
DB 3213 CACACCCGCGCGCTTAATGCGCGCTAACAGGGCGCGTCCCATTCGACATTCAGGCTGCG 480  
QY 481 CAACGTGTTGGAAGGCGCATCGGTGCGGCGCTTCCTGCTTAATGCGACGCTGGGAAAG 540  
DB 3153 CAACGTGTTGGAAGGCGCATCGGTGCGGCGCTTCCTGCTTAATGCGACGCTGGGAAAG 540  
QY 541 GGGATGTGTCGAAGCCGATTAAGTTGGGTAAACCGCAGGGTTCCTCCAGTCAAGAGTTG 600  
DB 3093 TCTGTGATTTTGTGCAAAATTCCTGTTGGCTTTCGCTAACGCGCTGAATGCGCGAAT 600  
QY 601 TAAACGACGCGCAGTGAAGCGCGCTGCTTCACTCAAGTTTGAACCCGTGAGAGACG 660  
DB 3033 CGGAATGTAAATGTGCGCACAGCGGCTAACCGCAATCCGCGGCGATCCACATCACTCT 660  
QY 661 GCAAGCTGCGGTCGAATGTGTTTAAACAGCGTATGAGCAGATGAAGTCTGACAC 720  
DB 2973 GCGAGTGAATGCGCGCAATCAAC-----GCGTGTGCGAGCTGAATACCC 2926  
QY 721 GCTCAGAAACGCGAGCTGAATTAACCTAGAAAGATATCATTTGTGACGTACGTAA 780  
DB 2925 GTTCAGAAATCGACTCTGATTTGCGGGTCACTTCGCGCACTAATGCGGTGCC 2865  
QY 781 AGATTAATCATGCGTAAATTTGACGATGGGATCTGTAATACGACTCACTAATAGGCGAAT 840  
DB 2865 AGCGGATAGCGTATGACCTGACGAGGTAGATCCGTTGGTGAAGTTTTCGATCTT 2806  
QY 841 TGGGTACCGGGCCCCCTCGAGGTGACGATATGATTAAGATTCGATTCGATTCCTGC 900  
DB 2805 TGGTATTAACAGGTTTCTGTCCGTAAAGCAAGCAAGATCGGATGTAATGTTCTT 2746  
QY 901 AGCCGGGGGATTCACATGTTTAAAGCGCGCGCACCGCGTGAAGCTCAAGCTTTTGT 960  
DB 2745 TGGCGAGCGGGTGGCGAGGTACCGGCGATCTCATATATGTCAGAGTTTAAACA 2686  
QY 961 TCCCTTTAGTGAAGGTTAATTAAGATCCATGCGCTAATTTTAAAGCACTAATCTTCTAG 1020  
DB 2685 GCTCCGAGAGATCTTTTCGTTATCGATGCGGAATG--CCCTGAAATAATCGGTTGC 2628

QY 1021 GGTAAATAGCTGATCAGATATATTCGTGGGCTTTTTCGGGCTCAAGTATCGCC 1080  
DB 2627 CACCCCCCTGGGCGAGGTGACAGATCTGCTGTGCTGTTTTCGGCGCTCGATTTGCG 2568  
QY 1081 CAAGTCGCGCTATCTGGGATCGGGAGAGAAACCCGTGCTTTTCCGCGAGTTG 1140  
DB 2567 CAGCATACGGGCTTCTCATACATTCGCTGATATTTAAACCTGAATATGCGCACTTC 2508  
QY 1141 AAGCG--CATGAAAGAGTTTCCGAGATGACTGTGCTGATTAAGGTTAGAGAAA 1199  
DB 2507 CCGGCGGTGGGCTGAAGTTTCAAGCTGTGATGACCTGTTCATTTTCAATATATA 2448  
QY 1200 CGACGTTTACATGATGATTCGGGAAGGTGGGCAATGACAGCCTTTAAAGCTGAACTG 1259  
DB 2447 ATCCGCTTGGTGGATTCGTTTCGCGGATTCGCTCATGGGTAAAGCCAGGGTT 2388  
QY 1260 TTGCTTCAAGGCCACCTGGGATTCAGATTCGTGCGGCTTTTCCGGAACAGTTCCGATG 1319  
DB 2387 TTCCAGTCAAGACGTTGTAACAGCCGCGAGTGAAGCGGCGTAAATACACTCACTATA 2328  
QY 1320 GTACGCGCGAAGCGGATCAGCAACCGAAACAATACCGGCAAGCGGAACTGCGCTGCC 1379  
DB 2327 G-----GCGAATTTGGGATCCGGGCCCCCTCGAGGTGACGCTATCGATTAAGCT 2277  
QY 1380 GGTGTGAGATTAATGACAGCGGTGCGCGCTGGGATTAATACGTACAGCAGAGCGGTTAT 1439  
DB 2276 TGAATATCAAAATTCGTGACGCCGGGGGATCACAATGTTTAAAGCGGCCGACCGGCT 2217  
QY 1440 CTTGCTGGATGCGCGAATAATGACATGATACCCCGTGAATTAACCGGCGCGGCT 1499  
DB 2216 GAGGCT-----CCAGCTTTTGTTCCTTTAGTGAAGGTTAATTTGGGCGCT 2172  
QY 1500 TGGCGTATCAATGCTCAATGCTTTCTGTGTGAATTTGTTATCGCTCAATTTGAC 1559  
DB 2171 TGGGCTAATCAATGCTCAATGCTTTCTGTGTGAATTTGTTATCGCTCAATTTGAC 2112  
QY 1560 ACAACATACAGCGCGAAGCATTAAGTTAAAGCCTGGGGTCCCTAATGATGAGTTAAC 1619  
DB 2111 ACAACATACAGCGCGAAGCATTAAGTTAAAGCCTGGGGTCCCTAATGATGAGTTAAC 2052  
QY 1620 TCAATTAATTTGCTTGGCTCACTGCGCGCTTTTCAGTGGGAAACCTGCTGCGCAG- 1678  
DB 2051 TCACATTAATTTGCTTGGCTCACTGCGCGCTTTTCAGTGGGAAACCTGCTGCGCAGC 1992  
QY 1679 -----CTGATTAATGATCGGCAAGCGCGGGGAGAGCGGTTTGGGTATGGGCG 1731  
DB 1991 TGCATGCTGATTAATGATGCGCAAGCGCGGGGAGAGCGGTTTGGGTATGGGCG 1932  
QY 1732 CTCTTCGCTTCTGCTGCTCACTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1791  
DB 1931 CTCTTCGCTTCTGCTGCTCACTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1872  
QY 1792 ATCAGCTCACTCAAGGCGGTAATACGTTATCCAGAGATTCAGGGATTAAGCAGAAA 1851  
DB 1871 ATCAGCTCACTCAAGGCGGTAATACGTTATCCAGAGATTCAGGGATTAAGCAGAAA 1812  
QY 1852 GAACATGAGCAAAAGCGCAAAAGCGCAAAAGCGCAAAAGCGCAAAAGCGCGCTTGTGCG 1911  
DB 1811 GAACATGAGCAAAAGCGCAAAAGCGCAAAAGCGCAAAAGCGCAAAAGCGCGCTTGTGCG 1752  
QY 1912 GTTTTTCATAGGCTCGGCCCCCTGACAGCATCAAAATTCGACGCTCAAGTCAGAG 1971  
DB 1751 GTTTTTCATAGGCTCGGCCCCCTGACAGCATCAAAATTCGACGCTCAAGTCAGAG 1692  
QY 1972 GTGCGAAACCGGACGATTAAGATTAAGATTAAGAGCGTTTCCCTGGAAGTCCCTGCT 2031  
DB 1691 GTGCGAAACCGGACGATTAAGATTAAGATTAAGAGCGTTTCCCTGGAAGTCCCTGCT 1632  
QY 2032 GCGCTCTCTGTTCCGACCTGCGGTTAACCGGAATCCTGTCGCTTTCCTTTCGGG 2091  
DB 1631 GCGCTCTCTGTTCCGACCTGCGGTTAACCGGAATCCTGTCGCTTTCCTTTCGGG 1572

Qy 2092 AAGGTGCGCTTCTCATAGTCAAGCTGATGATCTCAGTTGGATGATGCTTGG 2151  
Db 1571 AAGGTGCGCTTCTCATAGTCAAGCTGATGATCTCAGTTGGATGATGCTTGG 1512  
Qy 2152 CTCAAGCTGGGCTGTGTGCAAGAACCCCGTTAGCCGACCGCTGCTTATCCG 2211  
Db 1511 CTCAAGCTGGGCTGTGTGCAAGAACCCCGTTAGCCGACCGCTGCTTATCCG 1452  
Qy 2212 TAACTATGCTTGAAGTCAACCCGCTAAGACAGCATTTATCGCACTGGACAGCCAC 2271  
Db 1451 TAACTATGCTTGAAGTCAACCCGCTAAGACAGCATTTATCGCACTGGACAGCCAC 1392  
Qy 2272 TGGTAACAGATTAGCAGAGGATATGATGGGCTGCTACAGATCTTGAAGTGTG 2331  
Db 1391 TGGTAACAGATTAGCAGAGGATATGATGGGCTGCTACAGATCTTGAAGTGTG 1332  
Qy 2332 GCTTAACAGCTACACTAGAAAGACAGATTTGGTATCTGCGCTCTGTGAAGCCAGT 2391  
Db 1331 GCTTAACAGCTACACTAGAAAGACAGATTTGGTATCTGCGCTCTGTGAAGCCAGT 1272  
Qy 2392 TACCTTGGGAAAAAGATTGGTAGCTCTTGATCCGCAACAAACACCGCTGGTAGCG 2451  
Db 1271 TACCTTGGGAAAAAGATTGGTAGCTCTTGATCCGCAACAAACACCGCTGGTAGCG 1212  
Qy 2452 TGGTTTTTTTGTGGCAAGCAGATTAGCGCGAGAAAAAGGATCTCAAGAAATCC 2511  
Db 1211 TGGTTTTTTTGTGGCAAGCAGATTAGCGCGAGAAAAAGGATCTCAAGAAATCC 1152  
Qy 2512 TTTGATCTTTTCTACGGGGCTGACGCTCAGTGAACGAAAACTCACCTTAAGGATTT 2571  
Db 1151 TTTGATCTTTTCTACGGGGCTGACGCTCAGTGAACGAAAACTCACCTTAAGGATTT 1092  
Qy 2572 GGTATGAGATTATCAAAAAAGATCTTCACTAGATCTTTAAATTAATGAAGTT 2631  
Db 1091 GGTATGAGATTATCAAAAAAGATCTTCACTAGATCTTTAAATTAATGAAGTT 1032  
Qy 2632 TAAATCAATCTAAAGTATATATGATGAATCTGGCTGACGATTACCAATGCTTAATCAG 2691  
Db 1031 TAAATCAATCTAAAGTATATATGATGAATCTGGCTGACGATTACCAATGCTTAATCAG 972  
Qy 2692 TGAAGCACTATCTCAAGCATCTGTCTATTTGCTTACATCAGTATGCTGACTCCCGT 2751  
Db 971 TGAAGCACTATCTCAAGCATCTGTCTATTTGCTTACATCAGTATGCTGACTCCCGT 912  
Qy 2752 CGTGAATTAATCTACGATACGGAGGGCTTACATCTGGCCCCAGTGTGCAATGATACC 2811  
Db 911 CGTGAATTAATCTACGATACGGAGGGCTTACATCTGGCCCCAGTGTGCAATGATACC 852  
Qy 2812 GCGAGACCCAGCTCACCGGCTCAGATTTATCAGCAATTAACAGCCAGCCGAGAGGAC 2871  
Db 851 GCGAGACCCAGCTCACCGGCTCAGATTTATCAGCAATTAACAGCCAGCCGAGAGGAC 792  
Qy 2872 CGAGCGAGAAAGTGTCTGTGCACTTTATCCGCTCCATCCAGTCTATTAATTTGTCGG 2931  
Db 791 CGAGCGAGAAAGTGTCTGTGCACTTTATCCGCTCCATCCAGTCTATTAATTTGTCGG 732  
Qy 2932 GGAAGCTAAGTAAGTATGCTGCAAGTAAATGTTGGCAAGCTTGTGCAATGCTAC 2991  
Db 731 GGAAGCTAAGTAAGTATGCTGCAAGTAAATGTTGGCAAGCTTGTGCAATGCTAC 672  
Qy 2992 AGGATGCTGGTCAAGCTGCTGCTTGTGATGGCTTCAAGCTCCGCTTCCCAAGC 3051  
Db 671 AGGATGCTGGTCAAGCTGCTGCTTGTGATGGCTTCAAGCTCCGCTTCCCAAGC 612  
Qy 3052 ATCAAGCGATTACATGATCCCATGTTGTGCAAAAAAGCGGTAGCTCTCCGATCC 3111  
Db 611 ATCAAGCGATTACATGATCCCATGTTGTGCAAAAAAGCGGTAGCTCTCCGATCC 552  
Qy 3112 TCCGATGCTGTCAAGATGATGGCGCAGTGTATCACTCATGTTATGCAAGCACT 3171  
Db 551 TCCGATGCTGTCAAGATGATGGCGCAGTGTATCACTCATGTTATGCAAGCACT 492  
Qy 3172 GCATTAATCTCTTACTGTCAATGCAATCGTAAAGATGCTTTTCTGTGACTGATGATCTC 3231

Db 491 GCATTAATCTCTTACTGTCAATGCAATCGTAAAGATGCTTTTCTGTGACTGATGATCTC 432  
Qy 3332 AACCAAGCATTTAGAAATAGTATGATGCGCGACCGAGTTGCTTTGCCCGGCTCAAT 3291  
Db 431 AACCAAGCATTTAGAAATAGTATGATGCGCGACCGAGTTGCTTTGCCCGGCTCAAT 372  
Qy 3292 ACGGATTAATACCGCGCACATACAGAACTTTAAAGTGTCTCATTTGAAAAAGCTTC 3351  
Db 371 ACGGATTAATACCGCGCACATACAGAACTTTAAAGTGTCTCATTTGAAAAAGCTTC 312  
Qy 3352 TTGGGGGGAATACTCAAGGATCTTACCGCTGTAGATCCGTTGCAATGAATCCAC 3411  
Db 311 TTGGGGGGAATACTCAAGGATCTTACCGCTGTAGATCCGTTGCAATGAATCCAC 252  
Qy 3412 TCGTGAACCAACTGATCTTCAAGCATTTTACTTTTACCAAGCGTTTCTGGGTAGCAAA 3471  
Db 251 TCGTGAACCAACTGATCTTCAAGCATTTTACTTTTACCAAGCGTTTCTGGGTAGCAAA 192  
Qy 3472 AACGAAAGCAAAATGCGCAAAAAAGGAATAGGCGCACAGCAAAATGTTGAATACT 3531  
Db 191 AACGAAAGCAAAATGCGCAAAAAAGGAATAGGCGCACAGCAAAATGTTGAATACT 132  
Qy 3532 CATACCTCTCTTTTCAATTAATTTAGAGCATTTACAGGCTTATGCTCATGAGCGG 3591  
Db 131 CATACCTCTCTTTTCAATTAATTTAGAGCATTTATCAGGCTTATGCTCATGAGCGG 72  
Qy 3592 ATACATATTTGAATGATTTTGAAGAAATTAACAATAGGGGTTCCGCGCATTTCCCG 3651  
Db 71 ATACATATTTGAATGATTTTGAAGAAATTAACAATAGGGGTTCCGCGCATTTCCCG 12  
Qy 3652 AAAAGTCCAC 3662  
Db 11 AAAAGTCCAC 1

RESULT 13  
CQ768840/c  
LOCUS CQ768840 3918 bp DNA linear PAT 04-MAR-2004  
DEFINITION Sequence 59 from Patent WO2004007550.  
ACCESSION CQ768840  
VERSION CQ768840.1 GI:45112122  
KEYWORDS  
SOURCE  
ORGANISM  
synthetic construct  
synthetic construct  
artificial sequences.  
REFERENCE  
1 Unger, C.M., Zehetmeier, C., Torella, C., Niewoehner, J., Ahrens, B. and Beste, G.  
Mcam Inhibitors  
Patent: WO 2004007550-A 59 22-JAN-2004;  
Xerion Pharmaceuticals GmbH (DE)  
Location/Qualifiers  
1. 3918  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="Description of Artificial Sequence: plasmid"

ORIGIN  
Query Match 68.8%; Score 2517.8; DB 6; Length 3918;  
Best Local Similarity 81.8%; Pred. No. 0;  
Matches 3052; Conservative 0; Mismatches 537; Indels 142; Gaps 7;  
Qy 3 AAATGTAGCGTAAATATTTGTTAAATTCGCGTAAATTTTGTAAATCAGCTCAT 62  
Db 3732 AAATGTAGCGTAAATATTTGTTAAATTCGCGTAAATTTTGTAAATCAGCTCAT 3673  
Qy 63 TTTTAAACAATAGCGCAAAATCCGCAAAATCCCTTAATAATCAAAAGATAGACGAGA 122  
Db 3672 TTTTAAACAATAGCGCAAAATCCGCAAAATCCCTTAATAATCAAAAGATAGACGAGA 3613  
Qy 123 TAGGGTTAGTGTGTTCCAGTTTGAACAAGATCCACTTAATTAAGAACGTGACTCCA 182

```

Db      3612 TAGGGTGAATGATTTGTTCCAGTTTGGAAACAAGATCCACTATTAAGAAAGTGGACCTGA 3553
Qy      183 AGCTGAAGAGCGGAAAAAACCGTCTATCAGGGCGATGGGCCACTACGGAAACATACCTT 242
Db      3552 AGGTAAAGAGACGAAAAACCGTCTATCAGGGCGATGGGCCACTACGGAAACATACCTGA 3493
Qy      243 AATCAAGTTTTTTGGGGGTGAGAGTCCGCTAAAGCATCTAAATCCGAACCTTAAAGGAGCC 302
Db      3492 AATCAAGTTTTTTGGGGGTGAGAGTCCGCTAAAGCATCTAAATCCGAACCTTAAAGGAGCC 3433
Qy      303 CCGCATTTAAGCTTTGACGGGAAAGCCGGCGAAACGTGGCGAAGAAAGGAAAGAAAG 362
Db      3432 CCGCATTTAAGCTTTGACGGGAAAGCCGGCGAAACGTGGCGAAGAAAGGAAAGAAAG 3373
Qy      363 CGAAGAGAGCGGGCGCTAAGGGCGCTGGCAAGTGAAGCGTACCGCTGCGCTGAACCA 422
Db      3372 CGAAGAGAGCGGGCGCTAAGGGCGCTGGCAAGTGAAGCGTACCGCTGCGCTGAACCA 3313
Qy      423 CACCCGCGCGCTTAATGCGCGCTACAGGGCGCT----- 458
Db      3312 CACCCGCGCGCTTAATGCGCGCTACAGGGCGCGTACTATGCTTGAAGTATGCG 3253
Qy      459 ----- 471
Db      3252 GTGTGAATATACCCACAGATGCGTAAAGAAATATACGCATCAGGCGCAATGGCCATT 3193
Qy      472 CAGGCTGCGCAACTGTTGGGAAGGGCGATCGGTGGGGCGCTTTCGCTAATACCGCACT 531
Db      3192 CAGGCTGCGCAACTGTTGGGAAGGGCGATCGGTGGGGCGCTTTCGCTAATACCGCACT 3133
Qy      532 GGGGAAAGGGGATGTCGTCGCAAGGCGATTAAGTTGGGTAAAGCCAGGGTTTTCCAGTC 591
Db      3132 GGGGAAAGGGGATGTCGTCGCAAGGCGATTAAGTTGGGTAAAGCCAGGGTTTTCCAGTC 3073
Qy      592 ACAGAGTTGTAAACGACGCGCAAGTGAAGCGCGCCGCTGTTCAATACGTTTTGAACCGT 651
Db      3072 ACAGAGTTGTAAACGACGCGCAAGTGAAGCGCGCCGCTGTTCAATACGTTTTGAACCGT 3013
Qy      652 GGAGGACGGGCAACTCCGCGTCAATGTGTTTTTACAGAGTGAAGAGATGAAGAT 711
Db      3012 -----TAGAAACGTCGAAATATACATATAGTGGCAACATTAAGAAA 2964
Qy      712 GCTGCAACGCTGAGAACAGCAGCTAGATTAACCTAGAAAGATATATTTGTGAC 771
Db      2963 CGCAAGACACACGGAATTAAGTTATTTGTGCAATCAATGAAGAAATTCATATGTTT 2904
Qy      772 GTACGTTAAAGATATATATGCTTAATTAAGCATGGATCTGTATATGCACTCAT 831
Db      2903 ACCAGCGCCAAAGCATTAAGGGGACATTCACCGACTGAGGCAAGAAAGTAAATATG 2844
Qy      832 AGGCGAATGGGTACCGGGCCCCCTCGAGGTGCAAGGTATCGATAGCTTGAATTCG 891
Db      2843 ACCGAAATATATCTTAAGTGAATTTATCCGTCACGACTTGAAGCTTGAATTCG 891
Qy      892 AATTCCGCGACCCCGGGGATTCAGTATTTAGAGCGGCGCCACCGCGGTGAGCTCC 951
Db      2783 AGAGCCAGCA-----AATCACCAAGTACACCATTTACATTAAGCAAGCGGAAACG 2732
Qy      952 AGCTTTTGTCCCTTATAGTGGGTAAATTAAGATCCCATGCGCTCAATTTTACGAGACTA 1011
Db      2731 TCACCAATGAACATCATGATAGACGACGCTAATCATGAGCAAGATCA---AGTTTG 2675
Qy      1012 TCTTTTAAAGGTTAATCTAGTGCATCAGAGCATATATGTCGGGCTTTTTCGGGCTCA 1071
Db      2674 CTTTATGCTCAGACTGATGCGGTTTTCATCGGCAATTTTTCGTCATAGCCCCCTTAATTA 2615
Qy      1072 GTCATGCGCCAAGCTGAGCGCTATCTGGGATCTGGGAGAAAGAAAGCCGTCCTTTTCCC 1131
Db      2614 GGGTTTGCATTTTTCATATCAAAATCACCGGAACCGAGCGCCACCGGAACCGCCA 2555
Qy      1132 GGAAGTTGAAGCGGCATGAAAAGATTTTCCGAGAGTGAATGCTGCTGATTTGACGTTG 1191

```

---

```

Db      2554 CCTCAGAGCCGCAACCTCAGAACCGCAACGCTCAGAGCCGCACTCAGAGCCGCCA 2495
Qy      1192 AGCGAAAAACGAGCTTTTACATGATGATTTGGGAAGGTGTGGCATGACGCTTTAAACG 1251
Db      2494 CCABAACACAC-----CAGAGCCGCGCAACCTATCCGACGACGCGGTTCCAGGGAT 2440
Qy      1252 GTGAATGTTTCTTTCAGAGCCACTGGGATTAACAGTTTCGTGGGGCTTTTCCGGAACAGT 1311
Db      2439 CCGGATACGCGACACCGCGCACTGGCGCGCGTTGATGATGCACTTCGCGACCGCC 2380
Qy      1312 TCCGATGTCACCGCGGAAGCGCATGACAAACCCGAACATATCCGCGCAAGCCGGAAT 1371
Db      2379 AGAGCAACCTCGCCTGAAACCGCTTCACCTGTCA----- 2344
Qy      1372 GCCGTGCGGTGTGAGATTAATGAACAGCGGTGCGGCGTGGGATTAATACGTACAGAG 1431
Db      2343 -----CTGGAACCTGGGCAATGGCGGGTGGCGCATTAAGAAACAACTAAAGAA 2291
Qy      1432 ACGGATATCTGGCTGAGATGCGCGAATGGAATGAGATACCCGTGAGTTACCGGCG 1491
Db      2290 TTGCGAATATTAATTTTTCAGTTGAAATCTCAAAAAA-----GGCT 2244
Qy      1492 GGGCGCTTGGCGTATCATGATGATGCTGTTTCTGTGGAATTTTATCCGCTCAC 1551
Db      2243 CCAAGCTTGGGTAATCATGATGATGCTGTTTCTGTGGAATTTTATCCGCTCAC 2184
Qy      1552 AATTCACACAAACATACAGCGCGGAGACATTAAGTAAAGCTTGGGCTCTTAATGAT 1611
Db      2183 AATTCACACAAACATACAGCGCGGAGACATTAAGTAAAGCTTGGGCTCTTAATGAT 2124
Qy      1612 GAGCTAATCATTAATTTGCTGCTCACTGCGCGCTTCCAGTGGGAAACCTGTTC 1671
Db      2123 GAGCTAATCATTAATTTGCTGCTCACTGCGCGCTTCCAGTGGGAAACCTGTTC 2064
Qy      1672 GTGCGAGTGTATTAAGATTCGCGCAACGCGGGGAGAGCGGTTTCGATTTGGGCG 1731
Db      2063 GTGCGAGTGTATTAAGATTCGCGCAACGCGGGGAGAGCGGTTTCGATTTGGGCG 2004
Qy      1732 CTCTTCGCTTCTCGCTCACTGACTGCGTGGGCTGCTGCTTTCGCTGCGCGAGCGGT 1791
Db      2003 CTCTTCGCTTCTCGCTCACTGACTGCGTGGGCTGCTGCTTTCGCTGCGCGAGCGGT 1944
Qy      1792 ATCAGCTCATCAAGAGCGGTAATACGTTATCAACAGATCAGGGGATTAACGAGAAA 1851
Db      1943 ATCAGCTCATCAAGAGCGGTAATACGTTATCAACAGATCAGGGGATTAACGAGAAA 1884
Qy      1852 GAACATGTGACAAAAGGACAGCAAAAGGCGGAAACGTTAAAGGCTCGGTTGTGGC 1911
Db      1883 GAACATGTGACAAAAGGACAGCAAAAGGCGGAAACGTTAAAGGCTCGGTTGTGGC 1824
Qy      1912 GTTTTTCATAGGCTTCGCGCCCTGACGAGCATCAAAAATGCAAGCTTAATGCAAG 1971
Db      1823 GTTTTTCATAGGCTTCGCGCCCTGACGAGCATCAAAAATGCAAGCTTAATGCAAG 1764
Qy      1972 GTGGCGAAACCGGACAGGACTTAAGATACAGGCGTTTCCCTGGAAGCTCCCTGT 2031
Db      1763 GTGGCGAAACCGGACAGGACTTAAGATACAGGCGTTTCCCTGGAAGCTCCCTGT 1704
Qy      2032 GGGCTTCTCTGTTCCGACCTTGGCTTACCGGATACGTGTCCGCTTTCCTTCCGG 2091
Db      1703 GGGCTTCTCTGTTCCGACCTTGGCTTACCGGATACGTGTCCGCTTTCCTTCCGG 1644
Qy      2092 AAGCGTGGCGCTTCTCATATGCTCAAGCTTAAGATATCAAGTTGGGTGAGTGTGCG 2151
Db      1643 AAGCGTGGCGCTTCTCATATGCTCAAGCTTAAGATATCAAGTTGGGTGAGTGTGCG 1584
Qy      2152 CTCCAAGCTGGGCTGTGTGACGAACCCCGCTTACGCGCAACGCGTGGCGCTTAACCGG 2211
Db      1583 CTCCAAGCTGGGCTGTGTGACGAACCCCGCTTACGCGCAACGCGTGGCGCTTAACCGG 1524
Qy      2212 TAACTATGCTTGAAGTCAACCGGTAAGACAGACTTATGCGCATGGGACAGACAC 2271
Db      1523 TAACTATGCTTGAAGTCAACCGGTAAGACAGACTTATGCGCATGGGACAGACAC 1464

```



Oy 2272 TGGTAAACAGATTAAGCAGAGGAGGTATGTAGGCGGTCTACAGAGTTCTTGAAGTGTG 2331  
 Db 1463 TGGTAAACAGATTAAGCAGAGGAGGTATGTAGGCGGTCTACAGAGTTCTTGAAGTGTG 1404  
 Oy 2332 GCTTAATACAGGCTACACTAGAGGACAGATTTGGTATCTGCGCTCTGTGAAGCCAGT 2391  
 Db 1403 GCTTAATACAGGCTACACTAGAGGACAGATTTGGTATCTGCGCTCTGTGAAGCCAGT 1344  
 Oy 2392 TACCTTGGGAAAAAGATTGGTAGCTTTGATCCGGCAACAAACACCGCTGTGAGCGG 2451  
 Db 1343 TACCTTGGGAAAAAGATTGGTAGCTTTGATCCGGCAACAAACACCGCTGTGAGCGG 1284  
 Oy 2452 TGGTTTTTTTGGTGAAGCAGAGATTACCGCGGAGAAAAAGATCTCAAGAAAGATCC 2511  
 Db 1283 TGGTTTTTTTGGTGAAGCAGAGATTACCGCGGAGAAAAAGATCTCAAGAAAGATCC 1224  
 Oy 2512 TTTGATCTTTTCTACGCGGCTGACGCTCAGTGGAGCAAAAACTCAGCTTAAGGATTTT 2571  
 Db 1223 TTTGATCTTTTCTACGCGGCTGACGCTCAGTGGAGCAAAAACTCAGCTTAAGGATTTT 1164  
 Oy 2572 GGTCAATGATTTATCAAAAAAGATCTTCACTGATCTTTTAAATTAATAAGTAA 2631  
 Db 1163 GGTCAATGATTTATCAAAAAAGATCTTCACTGATCTTTTAAATTAATAAGTAA 1104  
 Oy 2632 TAAATCAATCTAAAGTATATATGATTAATCTTGTCTGACAGTTAACCAATCTTAATCAG 2691  
 Db 1103 TAAATCAATCTAAAGTATATATGATTAATCTTGTCTGACAGTTAACCAATCTTAATCAG 1044  
 Oy 2692 TGAGGCACTTATCTCAGCGATCTGTCTATTTGTTTCAATCAGTATGAGTCTGACTCCCGT 2751  
 Db 1043 TGAGGCACTTATCTCAGCGATCTGTCTATTTGTTTCAATCAGTATGAGTCTGACTCCCGT 984  
 Oy 2752 CGTGTAGATTAATCTACGATACGGAAGGAGCTTAACATCTGCGCCAGTGTGCAATGATAC 2811  
 Db 983 CGTGTAGATTAATCTACGATACGGAAGGAGCTTAACATCTGCGCCAGTGTGCAATGATAC 924  
 Oy 2812 GCGAGACCCAGCTACACCGGCTCCAGATTTATACAGCAATTAACAGCCAGCCGGAAGGCG 2871  
 Db 923 GCGAGACCCAGCTACACCGGCTCCAGATTTATACAGCAATTAACAGCCAGCCGGAAGGCG 864  
 Oy 2872 CGAGCGCAGAGTGGTCTGTGCACTTTATCCGCTCCATCCAGTCTATTAATTTGTTGCG 2931  
 Db 863 CGAGCGCAGAGTGGTCTGTGCACTTTATCCGCTCCATCCAGTCTATTAATTTGTTGCG 804  
 Oy 2932 GGAAGCTAAGATTAAGTATGCGCAGTTAATGTTTGGCAGACGTTGTGCCATTGCTAC 2991  
 Db 803 GGAAGCTAAGATTAAGTATGCGCAGTTAATGTTTGGCAGACGTTGTGCCATTGCTAC 744  
 Oy 2992 AGGATCGTGTGTCAAGCTGTGTGTGTATGAGCTTCAATCCGCTCCGATCCCAAG 3051  
 Db 743 AGGATCGTGTGTGTCAAGCTGTGTGTGTATGAGCTTCAATCCGCTCCGATCCCAAG 684  
 Oy 3052 ATCAAGCGAGATTAATGATCCCAATGTTGTGCAAAAAAGCGGTAGCTCTTCCGATCC 3111  
 Db 683 ATCAAGCGAGATTAATGATCCCAATGTTGTGCAAAAAAGCGGTAGCTCTTCCGATCC 624  
 Oy 3112 TCCGATCGTGTGTCAAGATTAAGTGGCGCAGTGTATCACTCAATGTTATGAGCAGCACT 3171  
 Db 623 TCCGATCGTGTGTCAAGATTAAGTGGCGCAGTGTATCACTCAATGTTATGAGCAGCACT 564  
 Oy 3172 GCAATATTTCTTACTGTATGCAATCCGTAAGATGCTTTTCTGTGACTGTGTGATCTC 3231  
 Db 563 GCAATATTTCTTACTGTATGCAATCCGTAAGATGCTTTTCTGTGACTGTGTGATCTC 504  
 Oy 3232 AACCAAGTCAATCTGAGAAATGATGTATGCGGAGCCAGATGCTCTTGCCCGGCTCAAT 3291  
 Db 503 AACCAAGTCAATCTGAGAAATGATGTATGCGGAGCCAGATGCTCTTGCCCGGCTCAAT 444  
 Oy 3292 ACGGATTAATACCGCGCCACATAGCAGAACTTTAAAGTGTCTCAATTTGGAACACGTTCC 3351  
 Db 443 ACGGATTAATACCGCGCCACATAGCAGAACTTTAAAGTGTCTCAATTTGGAACACGTTCC 384

Oy 3352 TTGGGGGGAAGAACTCTCAAGAGATCTTACCGCTGTGAGATCCAGTTGATGTAACCCAC 3411  
 Db 383 TTGGGGGGAAGAACTCTCAAGAGATCTTACCGCTGTGAGATCCAGTTGATGTAACCCAC 324  
 Oy 3412 TCGTGAACCCAACTGATCTTTCAGCATCTTTTACTTTACACGACGCTTTCTGGGTAGCAGA 3471  
 Db 323 TCGAGACCCAACTGATCTTTCAGCATCTTTTACTTTACACGACGCTTTCTGGGTAGCAGA 264  
 Oy 3472 AACGGAAGGCAAAATCCGCGCAAAAAAGGAAATAGGCGCAGCAGCAAAATGTTGAATCT 3531  
 Db 263 AACGGAAGGCAAAATCCGCGCAAAAAAGGAAATAGGCGCAGCAGCAAAATGTTGAATCT 204  
 Oy 3532 CATACCTCTCTTTTCAATTAATTTAGAGCATTTACAGGCTTATGTCATGAGAGCG 3591  
 Db 203 CATACCTCTCTTTTCAATTAATTTAGAGCATTTACAGGCTTATGTCATGAGAGCG 144  
 Oy 3592 ATACATATTTGAATGTATTTAGAAAAATTAACAATAGGAGTTCCGCGCATTTCCCG 3651  
 Db 143 ATACATATTTGAATGTATTTAGAAAAATTAACAATAGGAGTTCCGCGCATTTCCCG 84  
 Oy 3652 AAAAGTCCAC 3662  
 Db 83 AAAAGTCCAC 73  
 RESULT 14  
 CQ802124/c 3918 bp DNA linear PAT 05-MAY-2004  
 LOCUS CQ802124  
 DEFINITION Sequence 109 from Patent WO200400717.  
 ACCESSION CQ802124  
 VERSION CQ802124.1 GI:47058654  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 REFERENCE  
 1  
 AUTHORS Unger C.M., Beese G., Jensen K.H., Zehetmeier C., Knauer R.,  
 Guelbenzu B.L., Torella C. and Ilag L.U.  
 TITLE Proteomic screen to identify disease-related biological  
 JOURNAL Molecules and inhibitors thereto  
 PATENT Patent: WO 200400717-A 109 22-JAN-2004;  
 Xerion Pharmaceuticals AG (DE)  
 FEATURES  
 source  
 1..3918  
 /organism="synthetic construct"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:32630"  
 /note="Description of Artificial Sequence: vector pXp10"  
 ORIGIN  
 Query Match 68.8%; Score 2517.8; DB 6; Length 3918;  
 Best Local Similarity 81.8%; Pred. No. 0;  
 Matches 3052; Conservative 0; Mismatches 537; Indels 142; Gaps 7;  
 Oy 3 AAATGTAAACGTTAATATTTTGTAAATTCGCGTTAAATTTTGTAAATCAGCTCAT 62  
 Db 3732 AAATGTAAACGTTAATATTTTGTAAATTCGCGTTAAATTTTGTAAATCAGCTCAT 3673  
 Oy 63 TTTTAAACAATAGGCGCAAAATCCCTTAATCAAAAAGAAATAGCAGCA 122  
 Db 3672 TTTTAAACAATAGGCGCAAAATCCCTTAATCAAAAAGAAATAGCAGCA 3613  
 Oy 123 TAGGTTGAGTGTGTTTCCAGTTTGAACAAGAGTCCATTAATTAAGAAAGTGGACTCA 182  
 Db 3612 TAGGTTGAGTGTGTTTCCAGTTTGAACAAGAGTCCATTAATTAAGAAAGTGGACTCA 3553  
 Oy 183 ACGTCAAAAGGCGAAAAACGCTCTATCAAGGCGAATGCGCCATCACTGAAACCATCACT 242  
 Db 3552 ACGTCAAAAGGCGAAAAACGCTCTATCAAGGCGAATGCGCCATCACTGAAACCATCACT 3493  
 Oy 243 AATCAAGTTTGTGGGTGAGAGTGCCTTAAGCACTAAATCGAAACCTTAAGGAGGCC 302  
 Db 3492 AATCAAGTTTGTGGGTGAGAGTGCCTTAAGCACTAAATCGAAACCTTAAGGAGGCC 3433

QY 303 CCGATTAGACTTGAACGGGAAACCGGCGAAGCTGCGAGAAAGAAAGAAAG 362  
Db 3432 CCGATTAGACTTGAACGGGAAACCGGCGAAGCTGCGAGAAAGAAAGAAAG 3373  
QY 363 CCAAAAGAGCGGCGCTAGAGCGCTGAGCAAGTGAAGCGTCAAGCTGCGGTAA 422  
Db 3372 CCAAAAGAGCGGCGCTAGAGCGCTGAGCAAGTGAAGCGTCAAGCTGCGGTAA 3313  
QY 423 CACCGCGCGCTTAATGCGCGCTACAGGCGCGT----- 458  
Db 3312 CACCGCGCGCTTAATGCGCGCTACAGGCGCGCTTAATGCGCTTAATGCG 3253  
QY 459 -----CCCATTCGCATT 471  
Db 3252 GTGTAAATACCGCAAGATCGTTAAGAAATAACCGCATAGCGCATTCGCATT 3193  
QY 472 CAGGCTGCGCACTGTTGGGAAGGCGATGCGGCGCTCTTGCCTATTAGCGAGCT 531  
Db 3192 CAGGCTGCGCACTGTTGGGAAGGCGATGCGGCGCTCTTGCCTATTAGCGAGCT 3133  
QY 532 GCGCAAGGGGAGTGTCTCAAGGCGATTAAAGTGAAGCGCAAGGCTTTCCAGTC 591  
Db 3132 GCGCAAGGGGAGTGTCTCAAGGCGATTAAAGTGAAGCGCAAGGCTTTCCAGTC 3073  
QY 592 ACAGCGTTGTAACGACGCGCGAGTGAAGCGCGCTGTCATTCAAGTTTGAACCGGT 651  
Db 3072 ACAGCGTTGTAACGACGCGCGAGTGAAGCGCGCTGTCATTCAAGCTTAATG 3013  
QY 652 GAGAGACGCGCAGACTCGCGGTGCAAAATGTTTACAGCGTATGAGCAGATGAAGT 711  
Db 3012 -----TAGCAAACTGCAAAATCATACATTAAGGTGGCAATTAAGAA 2964  
QY 712 GCTCGACAGCTGCGCAACACGCACTAGATTAACTAGAAAGTATCATATTGAC 771  
Db 2963 CCGAAGACACGCAAGATTAAGTTATTGTCACAAATCATTAAGAAATCATTAAGTTT 2904  
QY 772 GATCGTTAAGATTAATCATGCGTAAATTAAGCATGAGATCTGTAATACGACTCAT 831  
Db 2903 ACCAGCGCCAAAGCATTAAGGCGCATTAACGCACTGAGGCAAAAGTAATATG 2844  
QY 832 AGGCGAATGGGTACCGGCGCCCCCTGAGGTGACGAGTATGATTAACCTTGATG 891  
Db 2843 ACGCAATTAATTAATTAAGGTGAATTAATACCGTCAACGACTTGAGCCATTGGGAAT 2784  
QY 892 AATTCTGACGCGCGGAGATCACTAGTCTAGAGCGCGCCACCGCGGTGAGCTCC 951  
Db 2783 AGAGCAGCA-----AAATCAGCAGTAGCACTTACATTAGCAAGCGGAAAG 2732  
QY 952 AGCTTTGTTCCCTTAAAGGTGAAGTGAATCCATGCGCATTTTACGAGACTA 1011  
Db 2731 TCACCAATGAACCATGATAGAGCAACGTAATCATGAGCAAGCAATCA--AGTTG 2675  
QY 1012 TCTTTAGAGGTAAATTAAGTCAATGATATGCGTGGGTCTTTTCCGCTCA 1071  
Db 2674 CTTTACGTCACACTGTAGCGCGTTTCAATGCGCAATTTTCCGTAATAGCCCTTAAT 2615  
QY 1072 GTATCGCCCAAGCTGCGCTATCTGAGCATCGGAGAGAAAGACCGGCTTTTCC 1131  
Db 2614 GCGTTTGCATTTTTCATTAATCAAAATCAACCGGAACCGGACCGCAACCGCA 2555  
QY 1132 GCGAGGTTGAAGCGCATGGAAGATTGCGAGAGATGACTGTCTGCAATTAAGCTG 1191  
Db 2554 CCGTACAGACCGGCAACCTCAAGAACCGCACTCAGAGCGGCAACCTCAGAGCCGCA 2495  
QY 1192 AGGAAAGACGAGTTTACATGATGATTCGGAAGGTGAGCATGAGACGCTTTAAG 1251  
Db 2494 CCAAGAACACAC-----CAGAGCGCGCCAGCTATAGCGCAACGCGTTCCAGCGAG 2440  
QY 1252 GTGAACGTGTTGTTGAGCGCACTTGGATACCAAGTTGTCGCGCTTTCCGAGCAG 1311  
Db 2439 CCGGATAGCGCACCGGCGACCTGCGCGCGCTTGTATGAGTCACTTCCGCAACGCC 2380

QY 1312 TCCGAGTGTGAGCCCGAAGCGCATGAGCAACCGAACAATACCGGCGACAGCCGAGCT 1371  
Db 2379 AGAGCAACTCCGCTGAACCGCTTCCACTGTGCA----- 2344  
QY 1372 GCGTGCCTGTGTGAGATTAATGACAGCGGTGCGCGCTGGAGATATTAAGTCAAGAG 1431  
Db 2343 -----CTGACCTGGCCATGCGCGCTGGCGCATTAAGAAAGCAACTAAGAA 2291  
QY 1432 ACGGTATCTGTGCTGAGTCCGCAAGAAATGACATGATACCCGTGAGTTACCGGCG 1491  
Db 2290 TTGGAAATATATTTTTCAGCTTAAGAAATCTCAAAAAA-----GGCT 2244  
QY 1492 GCGCGCTTGGCTAATCATGTGATAGCTGTTCTGTGTGAATTTGATCCGCTAC 1551  
Db 2243 CCAAGCTTGGCTAATCATGTGATAGCTGTTCTGTGTGAATTTGATCCGCTAC 2184  
QY 1552 AATTCAACACATCATGAGCGGAGCAATTAAGTGAAGCTGAGGTGCTTAAGT 1611  
Db 2183 AATTCAACACATCATGAGCGGAGCAATTAAGTGAAGCTGAGGTGCTTAAGT 2124  
QY 1612 GAGCTAATCAATTAATGCTGCTGCTCACTGCGCTTTCCAGTCCGGAACCTGTC 1671  
Db 2123 GAGCTAATCAATTAATGCTGCTGCTCACTGCGCTTTCCAGTCCGGAACCTGTC 2064  
QY 1672 GTGCGAGCTGATTAATTAATGCGCAACGCGCGGAGAGCGGTTGCGTATGAGCG 1731  
Db 2063 GTGCGAGCTGATTAATTAATGCGCAACGCGCGGAGAGAGCGGTTGCGTATGAGCG 2004  
QY 1732 CTCTTCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1791  
Db 2003 CTCTTCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1944  
QY 1792 ATACGCTCACTCAAAAGGAGTAATAGGTTATCCAGAAATCAGGAGTAAACGAGAA 1851  
Db 1943 ATACGCTCACTCAAAAGGAGTAATAGGTTATCCAGAAATCAGGAGTAAACGAGAA 1884  
QY 1852 GAACTGTGAGCAAAAGCGCAGCAAAAGCGCAAGAACGTAAGAAAGCGGCTGTGCG 1911  
Db 1883 GAACTGTGAGCAAAAGCGCAGCAAAAGCGCAAGAACGTAAGAAAGCGGCTGTGCG 1824  
QY 1912 GTTTTTCATAGGCTCGCCCCCTGAGGATCAAGATCAAAATTCAGCTCAAGTCAAG 1971  
Db 1823 GTTTTTCATAGGCTCGCCCCCTGAGGATCAAGATCAAAATTCAGCTCAAGTCAAG 1764  
QY 1972 GTGCGCAACCGCAGACAGATTAAGATACAGGAGTTCGCCGTGAAGTCCCTGT 2031  
Db 1763 GTGCGCAACCGCAGACAGATTAAGATACAGGAGTTCGCCGTGAAGTCCCTGT 1704  
QY 2032 GCGCTCCCTGTTCCGACCCCTGCGCTTACCGATACCTGTCCGCTTCCCTGCGG 2091  
Db 1703 GCGCTCCCTGTTCCGACCCCTGCGCTTACCGATACCTGTCCGCTTCCCTGCGG 1644  
QY 2092 AAGCGTGGCTTTCATAGCTCAGCGTGAAGTATCAGTTCGCTGCTGCTGCTGCT 2151  
Db 1643 AAGCGTGGCTTTCATAGCTCAGCGTGAAGTATCAGTTCGCTGCTGCTGCTGCT 1584  
QY 2152 CTCGAAGCGGCTGTGTGACGAACCCCGCTTCAAGCCGACCGGCTGCTTATCCG 2211  
Db 1583 CTCGAAGCGGCTGTGTGACGAACCCCGCTTCAAGCCGACCGGCTGCTTATCCG 1524  
QY 2212 TAATATGCTCTTGAATCAACCGGATTAAGACAGATTAATGCTGCTGCTGCTGCT 2271  
Db 1523 TAATATGCTCTTGAATCAACCGGATTAAGACAGATTAATGCTGCTGCTGCTGCT 1464  
QY 2272 TGATTAAGAGTTTGAAGAGAGAGTATAGGCGGTCTACAGATTTCTTAAGAGTGT 2331  
Db 1463 TGATTAAGAGTTTGAAGAGAGAGTATAGGCGGTCTACAGATTTCTTAAGAGTGT 1404  
QY 2332 GCTTAATCAAGCTACACTGAAGACAGATTTGATTTGCTGCTGCTGCTGCTGCT 2391  
Db 1403 GCTTAATCAAGCTACACTGAAGACAGATTTGATTTGCTGCTGCTGCTGCTGCT 1344  
QY 2392 TACCTTGGAAGAAAGATTGTAGCTTGTATCCGCAAAACCAACGCTGTAGGCG 2451

Db 1343 TACCTCGGAAAAAGAGTTGGTAGCTTGTATCCGGCAAAACAACCGCTGGTAGAGCG 1284  
 Qy 2452 TGGTTTTTTTGTGGCAAGCAGCATTAACGGCGAAGAAAAAGAGATCTCAAGAGATCC 2511  
 Db 1283 TGGTTTTTTTGTGGCAAGCAGCATTAACGGCGAAGAAAAAGAGATCTCAAGAGATCC 1224  
 Qy 2512 TTGTATCTTTCTACGGGGGCTGACGGCTCAGTGGAGAAAGCAAGCTTAAGGGATTTT 2571  
 Db 1223 TTGTATCTTTCTACGGGGGCTGACGGCTCAGTGGAGAAAGCAAGCTTAAGGGATTTT 1164  
 Qy 2572 GGTATGATGATTAACAAAAAGATCTTACCTAGATCTTTTAAATTAATAAGAGTTT 2631  
 Db 1163 GGTATGATGATTAACAAAAAGATCTTACCTAGATCTTTTAAATTAATAAGAGTTT 1104  
 Qy 2632 TAAATCAATCTAAAGATATATAGATTAAGTTGGTGTGACAGTTTCAAGCTTAATCAG 2691  
 Db 1103 TAAATCAATCTAAAGATATATAGATTAAGTTGGTGTGACAGTTTCAAGCTTAATCAG 1044  
 Qy 2692 TGAGGACACCTATCTCAGGAGCTGTGTATTCGTTCAATCCATAGTGGCTGACCTCCCGT 2751  
 Db 1043 TGAGGACACCTATCTCAGGAGCTGTGTATTCGTTCAATCCATAGTGGCTGACCTCCCGT 984  
 Qy 2752 CGTGTAGATTAACGATACCGGAGGGCTTACCATCTGGCCCCAGTGTGCAATGATATAC 2811  
 Db 983 CGTGTAGATTAACGATACCGGAGGGCTTACCATCTGGCCCCAGTGTGCAATGATATAC 924  
 Qy 2812 GCGAGACCCACGCTTCACGGCTCCAGATTTATCAGCAATAAACGACGCGGAGAGGGC 2871  
 Db 923 GCGAGACCCACGCTTCACGGCTCCAGATTTATCAGCAATAAACGACGCGGAGAGGGC 864  
 Qy 2872 CGAGCGGAGAGTGGTCTCGCACTTATCGGCTCCATCCAGTCTAATTAATGTTCGCG 2931  
 Db 863 CGAGCGGAGAGTGGTCTCGCACTTATCGGCTCCATCCAGTCTAATTAATGTTCGCG 804  
 Qy 2932 GGAAGCTAGATAGTATAGTTGGCCAGTAAATAGTTGGCAACGTTGGTCCATGTATC 2991  
 Db 803 GGAAGCTAGATAGTATAGTTGGCCAGTAAATAGTTGGCAACGTTGGTCCATGTATC 744  
 Qy 2992 AGGATCGTGGTGTCAAGCTCGTGTGGTATGAGCTTCAATCAGCTCCGGTCCCAACG 3051  
 Db 743 AGGATCGTGGTGTCAAGCTCGTGTGGTATGAGCTTCAATCAGCTCCGGTCCCAACG 684  
 Qy 3052 ATCAAGCGAGTTCAGATATCCCAATGTTGTGCAAAAAAGCGGTAGCTCTTCGGTCC 3111  
 Db 683 ATCAAGCGAGTTCAGATATCCCAATGTTGTGCAAAAAAGCGGTAGCTCTTCGGTCC 624  
 Qy 3112 TCCGATCGTGTGTCAAGATAGTGGCCGAGTGTATCACTCATGTTATGGCAGCACT 3171  
 Db 623 TCCGATCGTGTGTCAAGATAGTGGCCGAGTGTATCACTCATGTTATGGCAGCACT 564  
 Qy 3172 GCATTAATCTCTTACTGTATGATCCATCGTAAGATGCTTTTCTGTGATGTGTGATCTC 3231  
 Db 563 GCATTAATCTCTTACTGTATGATCCATCGTAAGATGCTTTTCTGTGATGTGTGATCTC 504  
 Qy 3232 AACCAAGCATCTTGAAGATGTATGCGGCGAAGTGGCTCTTGGCCGGCTCAAT 3291  
 Db 503 AACCAAGCATCTTGAAGATGTATGCGGCGAAGTGGCTCTTGGCCGGCTCAAT 444  
 Qy 3292 ACGGATTAATACCGCGCACATAGCAGAACTTTAAAGTGTCTCATTTGAAAGAGTTC 3351  
 Db 443 ACGGATTAATACCGCGCACATAGCAGAACTTTAAAGTGTCTCATTTGAAAGAGTTC 384  
 Qy 3352 TTGGGGGCGAAAACTTCAAGAGATCTTACGGCTGTGATGATCCAGTTGATGATACCCAC 3411  
 Db 383 TTGGGGGCGAAAACTTCAAGAGATCTTACGGCTGTGATGATCCAGTTGATGATACCCAC 324  
 Qy 3412 TCGGACACCCAACTGATCTTCAAGATCTTTTACCTTCAACGAGCTTTCGGGTAGAGAA 3471  
 Db 323 TCGGACACCCAACTGATCTTCAAGATCTTTTACCTTCAACGAGCTTTCGGGTAGAGAA 264  
 Qy 3472 AACGAGAGCGAAATGCGCGCAAAAGGAAATAGGCGACAGGAAATGTTAAATCT 3531

Db 263 AACAGAGGCAAAATGCGCGCAAAAAAGGAAATAGGCGCGACAGGAAATGTGAAATCT 204  
 Qy 3532 CATACTCTTCTTTTCAATATTTATGAGCATTTATCAGGCTTATTTGTCTCATAGAGCG 3591  
 Db 203 CATACTCTTCTTTTCAATATTTATGAGCATTTATCAGGCTTATTTGTCTCATAGAGCG 144  
 Qy 3592 ATACATATTTGAATGTATTTAGAAAAATTAACAAATAGGGGTTCCGGCACATTTCCCG 3651  
 Db 143 ATACATATTTGAATGTATTTAGAAAAATTAACAAATAGGGGTTCCGGCACATTTCCCG 84  
 Qy 3652 AAAAGTGCAC 3662  
 Db 83 AAAAGTGCAC 73

RESULT 15  
 XXU35129 3886 bp DNA circular SYN 26-SEP-1995  
 LOCUS  
 DEFINITION Plasmid pBSL141 cloning vector, complete sequence.  
 ACCESSION U35129  
 VERSION U35129.1 GI:984901  
 KEYWORDS

SOURCE  
 ORGANISM  
 Cloning vector pBSL141  
 Cloning vector pBSL141  
 artificial sequences; vectors.

REFERENCE  
 AUTHORS Alexeyev,M.F., Shokolenko,I.N. and Croughan,T.P.  
 TITLE Improved antibiotic-resistance gene cassettes and omega elements for Escherichia coli vector construction and in vitro deletion/insertion mutagenesis

JOURNAL  
 MEDLINE Gene 160 (1), 63-67 (1995)  
 PUBMED 95354958  
 7628718

REFERENCE  
 PubMed 2 (bases 1 to 3886)

AUTHORS Hengen,P.N.  
 TITLE Direct Submission  
 JOURNAL Submitted (31-AUG-1995) Paul N. Hengen, Laboratory of Mathematical Biology, National Cancer Institute, Frederick, MD 21702-1201, USA

FEATURES  
 source  
 Location/Qualifiers  
 1..3886

CDS  
 /organism="Cloning vector pBSL141"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:42702"  
 /plasmid="pBSL141"  
 complement (806..1339)  
 /codon\_start=1  
 /transl\_table=1  
 /product="gentamicin 3'-acetyltransferase"  
 /protein\_id="AAC53614.1"  
 /db\_xref="GI:984902"

CDS  
 /organism="MUSSSNDVTQOGRPTKLGSSWGIIRTCRLGPDQVKSRAAL  
 DLGREGDVAATYOHQPSDYLGNLRSFTFLAALPDEAVGALAAVLPFEOP  
 RSBIIYIDLAVSGEHRROCIATLILNKHANALGAVIYVQADYDDBDAVALYTKL  
 GIREVWHPIDISTAT"  
 complement (2898..3758)  
 /EC\_number="3.5.2.6"  
 /function="ampicillin resistance"  
 /codon\_start=1  
 /transl\_table=1  
 /product="beta-lactamase"  
 /protein\_id="AAC53615.1"  
 /db\_xref="GI:984903"

# ORIGIN

Query Match 68.7%; Score 2514.6; DB 12; Length 3886;  
 Best Local Similarity 80.8%; Pred. No. 0;  
 Matches 3142; Conservative 0; Mismatches 519; Indels 226; Gaps 9;

QY 1 CTAATTTGTAAGCGTTAATATTTTGTAAATTCGCTAAATTTTGTAAATCAGCTC 60  
 Db 1 CTAATTTGTAAGCGTTAATATTTTGTAAATTCGCTAAATTTTGTAAATCAGCTC 60  
 QY 61 ATTTTAAACCAATAGCCGAAATCGGGAAATCCCTTAATATCAAAAGATAGACGA 120  
 Db 61 ATTTTAAACCAATAGCCGAAATCGGGAAATCCCTTAATATCAAAAGATAGACGA 120  
 QY 121 GATAGGTTGAGTGTGTTTCCAGTTTGGACACAGAGTCCACTATTAAAGAACGTGACTC 180  
 Db 121 GATAGGTTGAGTGTGTTTCCAGTTTGGACACAGAGTCCACTATTAAAGAACGTGACTC 180  
 QY 181 CAAGCTCAAAAGGCGGAAACCGCTCTATCAGGGGGAATGGCCCACTACGTAAACCATCACC 240  
 Db 181 CAAGCTCAAAAGGCGGAAACCGCTCTATCAGGGGGAATGGCCCACTACGTAAACCATCACC 240  
 QY 241 CTAATCAAGTTTGGGCTGAGAGTCCGTAAAGCACTAAATCGAACCTTAAGGGAG 300  
 Db 241 CTAATCAAGTTTGGGCTGAGAGTCCGTAAAGCACTAAATCGAACCTTAAGGGAG 300  
 QY 301 CCCCAGATTAGAGCTTGAAGGGGAAAGCCGGGAAACGTGGGAAAGGAGGAG 360  
 Db 301 CCCCAGATTAGAGCTTGAAGGGGAAAGCCGGGAAACGTGGGAAAGGAGGAG 360  
 QY 361 AGCGAAAGAGCGGGCGCTAGGGCGCTGGCAAGTGAAGCGTCAAGCTGCGGTAAC 420  
 Db 361 AGCGAAAGAGCGGGCGCTAGGGCGCTGGCAAGTGAAGCGTCAAGCTGCGGTAAC 420  
 QY 421 CACACCCGCGCGCTTAATGCGCGCTACAGGGCCGCTCCCACTTGGCATTCAGGCTGCG 480  
 Db 421 CACACCCGCGCGCTTAATGCGCGCTACAGGGCCGCTCCCACTTGGCATTCAGGCTGCG 480  
 QY 481 CAACCTTGGGAAAGGCGGATCGGTGCGGGCCCTCTCTGCTATTAAAGCCAGCTGGGAGAG 540  
 Db 481 CAACCTTGGGAAAGGCGGATCGGTGCGGGCCCTCTCTGCTATTAAAGCCAGCTGGGAGAG 540  
 QY 541 GGGATGTGCTGCAAGGCGATTAAGTTGGGTAAAGCCAGGGTTTCCAGTCAAGCAGCTTG 600  
 Db 541 GGGATGTGCTGCAAGGCGATTAAGTTGGGTAAAGCCAGGGTTTCCAGTCAAGCAGCTTG 600  
 QY 601 TAAACGACGCGCACTAGCGCGCTGCTGCTC----- 631  
 Db 601 TAAACGACGCGCACTAGCGCGCTGCTGCTC----- 631  
 QY 632 -----ATTCAAGTTTGAACCCGTGAGAGACGGGCA 663  
 Db 632 -----ATTCAAGTTTGAACCCGTGAGAGACGGGCA 663  
 QY 663 GACCTCGCGTCCAAATGTGTTTACAGCGTGAAGAGAGAGATGAGATGC-TGACACGCG 720  
 Db 663 GACCTCGCGTCCAAATGTGTTTACAGCGTGAAGAGAGAGATGAGATGC-TGACACGCG 720  
 QY 721 ATCCACTAGTTCTAGACCGCGCCGACCGCGGTGGAGCTCGAATTGACCGCTGCGCGAGA 780  
 Db 721 ATCCACTAGTTCTAGACCGCGCCGACCGCGGTGGAGCTCGAATTGACCGCTGCGCGAGA 780  
 QY 781 AGCGGATCTCGGCTTGAACGAAATTGTAGGTGCGGTACTTGGGTGCAATATCAAGAGTCA 840  
 Db 781 AGCGGATCTCGGCTTGAACGAAATTGTAGGTGCGGTACTTGGGTGCAATATCAAGAGTCA 840  
 QY 841 TCACTCTTCCGTATGCCCCAATTGTATAGAGAGCCACTGCGGAGATGCTACCGTAAT 900  
 Db 841 TCACTCTTCCGTATGCCCCAATTGTATAGAGAGCCACTGCGGAGATGCTACCGTAAT 900  
 QY 880 ACGACTCACTATAGGCGGAAATGGGTGACCGGCGC-----CCCCCTGAGAGTGAAGGTATGCA 877  
 Db 880 ACGACTCACTATAGGCGGAAATGGGTGACCGGCGC-----CCCCCTGAGAGTGAAGGTATGCA 877  
 QY 901 CTGCTTGAAGTATGATCAATTAAGCAACAAAGCGCTTGGCTCATGTGAGAGATGGA 960  
 Db 901 CTGCTTGAAGTATGATCAATTAAGCAACAAAGCGCTTGGCTCATGTGAGAGATGGA 960  
 QY 961 TGAAGCGGTGCAATGCGCTGCTGCGGTGCTGCGCGAGAACTGCGAGATCAATAGATAT 1020  
 Db 961 TGAAGCGGTGCAATGCGCTGCTGCGGTGCTGCGCGAGAACTGCGAGATCAATAGATAT 1020  
 QY 1021 AGATCTCACTAGCGGCTGCTCAAACTGGGCGAGAACTGAAGCCGAGAGCGCCAAACGA 1080  
 Db 1021 AGATCTCACTAGCGGCTGCTCAAACTGGGCGAGAACTGAAGCCGAGAGCGCCAAACGA 1080

QY 913 CCACTAGTTCTAGAGCGGCGCCCAACCGCGGTGAGACTCCAGCTTTGTTCCCTTAGTGA 972  
 Db 1081 CCGCTTCTTGGTCCAGAGCAGCAAGCGCATGATGATCTTACTACGAGCAAGTCCCGA 1140  
 QY 973 GGGTTAATAGATCCATGCGTCAATTTTACGAGACTATCTTTCAAGGTTAATCTTAC 1032  
 Db 1141 GGTATGCGAGTCCGCGTGAATGTGGAGATAGGTGCTACGCTCCGAACTACAGACGA 1200  
 QY 1033 TGCATCAGATCAATATGTCGGGCTTTTTCGGCTCAATGATGCGCAAGCTGCGCT 1092  
 Db 1201 AAAGATGAAGAGACCCCGCATGATTAATCTTGCACAGGGCCGACCTTACATGCGAA 1260  
 QY 1093 ATCTGGCATCGGGAGAGAAAGCCCGTTCCTTTCCCGGAGGATGAAGCGGATGA 1152  
 Db 1261 TGAATGCCATATCTTGAAGCACTTAACCTTTGTTTGAAGGAGACTGCTGCTGAACAT 1320  
 QY 1153 AAGAGTT--TGCGAGAGATGACTGCTGCTGCTTGAAGGTAAGGAAACGACGTTTAC 1210  
 Db 1321 CGTTGCTGCTGCTAATCGTTGCTGCTCA-TAACTCAAAATCATGACCCAGGGGTA 1379  
 QY 1211 CATGATGATTCGGGAAAGTGTGGCCATGACAGCGCTTTAAGGTGAATCTTCTTCAAGC 1270  
 Db 1380 ACGCGCTTCTGCTTGAATGCGCGAGATGATGATGATCAAAAAACATCATTAACAGC 1439  
 QY 1271 CACTGGGATACAGTTGCTGCGGCTTTTCCGACACAGTTTCGATGCTCAGCCGAA 1330  
 Db 1440 CATGAAACCCGCTACGCTGCGCTTACACAGCTGCTGCTGCTCAAGGTTTGAACAGTT 1499  
 QY 1331 GCGCATCAGCAACCGCAATACCGGCGACAGCGGAACTGCGCGGTGCTGA--- 1387  
 Db 1500 GCGTGAAGCGATACGCTACTTGCATTAACATTACCAACGAAACAGCTTAATGCTCAATTC 1559  
 QY 1388 GATTATAGACGCGTCCGCGCTGGGATTAATGCTGACGAGAGACGGGTATCTGCTG 1447  
 Db 1560 GAGATTAAGCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1619  
 QY 1448 GATCCCGCAAAATGA----- 1464  
 Db 1620 GATATCAATTCGCAAGCCCGGGGATCACTAGTTCTAGAGCGGCGCCACCGCGGTG 1679  
 QY 1465 -----CATGGAATACCCCGTGAATACCGGCGGCGCGCTTGGGATATCAATGCTC 1515  
 Db 1680 GAGCTCAGCTTTGTTCCCTTGAATGAGGATTAATGCGGCTTGGGATATCAATGCTC 1729  
 QY 1516 ATAGCTTCTCTGTGTAATTTGATTCGCTCAATTCACGAAATATACGAGCCGG 1575  
 Db 1740 ATAGCTTCTCTGTGTAATTTGATTCGCTCAATTCACGAAATATACGAGCCGG 1799  
 QY 1576 AAGCATTAAGTGAAGCCGTGGGCTGCTTAATGAGTGAAGTCACTAATATGCTT 1635  
 Db 1800 AAGCATTAAGTGAAGCCGTGGGCTGCTTAATGAGTGAAGTCACTAATATGCTT 1859  
 QY 1636 GCGCTCACTGCGCGCTTTCAGTGGGAAACCTGTCGACAGTGTGATTAATGATCGG 1695  
 Db 1860 GCGCTCACTGCGCGCTTTCAGTGGGAAACCTGTCGACAGTGTGATTAATGATCGG 1919  
 QY 1696 CCAAGCGCGGGGAGAGCGGCTTTCGATTTGAGCGCTTTCGCTTCCGCTCACTGA 1755  
 Db 1920 CCAAGCGCGGGGAGAGCGGCTTTCGATTTGAGCGCTTTCGCTTCCGCTCACTGA 1979  
 QY 1756 CTGCTGCTGCTGCTGCTGCTGCTGCGGAGAGGATTAAGTCACTCAAGCGGTAAT 1815  
 Db 1980 CTGCTGCTGCTGCTGCTGCTGCGGAGAGGATTAAGTCACTCAAGCGGTAAT 2039  
 QY 1816 ACGGTTATCCACAGATCAAGGAGTAAACGAGAGAAACATGTGAGCAAAAGCGCAGCA 1875  
 Db 2040 ACGGTTATCCACAGATCAAGGAGTAAACGAGAGAAACATGTGAGCAAAAGCGCAGCA 2099  
 QY 1876 AAAGGCCAGAACCGTAAAGGCGGCTGCTGCTGCTTTCATTAAGCTCCTGCTGCTGCT 1935  
 Db 2100 AAAGGCCAGAACCGTAAAGGCGGCTGCTGCTGCTTTCATTAAGCTCCTGCTGCTGCT 2159  
 QY 1936 TGAAGACATCAAAAATGACGCTCAAGTCAAGGTGCGAGAACCCGACGAGACTATA 1995

```

Db      2160 TGACGAGATCACAAAAATCGACGCTCAAGTCAGAGTGGGAAACCGGACGAGACTATA 2219
Qy      1996 AAGATACAGGCGTTTCCCTCGAAGCTCCCTCGTGGCTCTCTCTGTTCGACCCCTCC 2055
Db      2220 AAGATACAGGCGTTTCCCTCGAAGCTCCCTCGTGGCTCTCTCTGTTCGACCCCTCC 2279
Qy      2056 GCTTACCGGATACCTGTCCGCTTCTCCCTCGGGAAGCGTGGGCTTTCTCATAGTTC 2115
Db      2280 GCTTACCGGATACCTGTCCGCTTCTCCCTCGGGAAGCGTGGGCTTTCTCATAGTTC 2339
Qy      2116 ACGCTAGATATCTCAGTTCGAGTGTAGTGTGCTGCTCAAGCTGGGCTGTGTGACGA 2175
Db      2340 ACGCTAGATATCTCAGTTCGAGTGTAGTGTGCTGCTCAAGCTGGGCTGTGTGACGA 2399
Qy      2176 ACCCCCGCTTCAAGCCGACCGCTGCGCTTATCCGGTAAGTATGCTTGAAGTCCACCC 2235
Db      2400 ACCCCCGCTTCAAGCCGACCGCTGCGCTTATCCGGTAAGTATGCTTGAAGTCCACCC 2459
Qy      2236 GGTAAAGACAGACTTATGCGACCTGCGACGACCACTGCTAAACAGGATTTACAGAGCGAG 2295
Db      2460 GGTAAAGACAGACTTATGCGACCTGCGACGACCACTGCTAAACAGGATTTACAGAGCGAG 2519
Qy      2296 GTATGTAGGCGGTGCTACAGAGTCTTGAAGTGTGCGCTAACTACGCGCTACACTAGAG 2355
Db      2520 GTATGTAGGCGGTGCTACAGAGTCTTGAAGTGTGCGCTAACTACGCGCTACACTAGAG 2579
Qy      2356 GACAGTATTTGTATCTGCGCTCTGCTGAGCCAGTTACCTTCGGAAGAAAGTGTGTAG 2415
Db      2580 GACAGTATTTGTATCTGCGCTCTGCTGAGCCAGTTACCTTCGGAAGAAAGTGTGTAG 2639
Qy      2416 CTCTTGATCCGCGCAAAACCAACGCGCTGTAGCGGTGTGTGTGTGTGTGTGTGTGTGT 2475
Db      2640 CTCTTGATCCGCGCAAAACCAACGCGCTGTAGCGGTGTGTGTGTGTGTGTGTGTGTGT 2699
Qy      2476 GATTACGCGCAGAAAAAGAGTCTCAGAAAGATCTTGTATCTTTTCTACGCGGCTGCA 2535
Db      2700 GATTACGCGCAGAAAAAGAGTCTCAGAAAGATCTTGTATCTTTTCTACGCGGCTGCA 2759
Qy      2536 CGCTCAGTGAACGAAAACGTCAGCTTAAAGGATTTGTGTATGAGATTTACAAAAGGAT 2595
Db      2760 CGCTCAGTGAACGAAAACGTCAGCTTAAAGGATTTGTGTATGAGATTTACAAAAGGAT 2819
Qy      2596 CTTTACCTAGATCTTTTAAATTAATAAGTTTAAATCAATCTTAAAGTATTAATGA 2655
Db      2820 CTTTACCTAGATCTTTTAAATTAATAAGTTTAAATCAATCTTAAAGTATTAATGA 2879
Qy      2656 GTAAACTGTGCTGACAGTACCAATGCTTATCAGTAGAGGACCTATCTCAGGAGTCTG 2715
Db      2880 GTAAACTGTGCTGACAGTACCAATGCTTATCAGTAGAGGACCTATCTCAGGAGTCTG 2939
Qy      2716 TCTATTTGCTCATCTCAGTATGCTGACTCCCGCTGCTGATGATTAATCAATCAAGGA 2775
Db      2940 TCTATTTGCTCATCTCAGTATGCTGACTCCCGCTGCTGATGATTAATCAATCAAGGA 2999
Qy      2776 GGGCTTACATCTGGCCCCAGTGTGCTGCAATGATACCGGAGACCCACGCTCACCGGCTCC 2835
Db      3000 GGGCTTACATCTGGCCCCAGTGTGCTGCAATGATACCGGAGACCCACGCTCACCGGCTCC 3059
Qy      2836 AGATTTATCAGCAATTAACAGACGACGGAAGGGGCGAGCGCAGAAAGTGGCTCGCAAC 2895
Db      3060 AGATTTATCAGCAATTAACAGACGACGGAAGGGGCGAGCGCAGAAAGTGGCTCGCAAC 3119
Qy      2896 TTTATCCGCTTCAATCCAGTCTATTAATTTGTGCGGGAAGCTAGAGTAAAGTTCGCC 2955
Db      3120 TTTATCCGCTTCAATCCAGTCTATTAATTTGTGCGGGAAGCTAGAGTAAAGTTCGCC 3179
Qy      2956 AGTTAATAGTTGGGCAAGCTTGTGCTATGCTACAGGCAATCGTGTGTCAACGCTCGTC 3015
Db      3180 AGTTAATAGTTGGGCAAGCTTGTGCTATGCTACAGGCAATCGTGTGTCAACGCTCGTC 3239
Qy      3016 GTTGTGTAGGCTTCAATCAGCTCGGTTCCCAAGATCAAGGGAGTTACATGATCCCC 3075

```

```

Db      3240 GTTGTGTAGGCTTCAATCAGCTCGGTTCCCAAGATCAAGGGAGTTACATGATCCCC 3299
Qy      3076 CATGTGTGCAAAAAGCGGTAGCTCTTCTGCTCTCCGATCGTTGTCAAGATTAAGTT 3135
Db      3300 CATGTGTGCAAAAAGCGGTAGCTCTTCTGCTCTCCGATCGTTGTCAAGATTAAGTT 3359
Qy      3136 GGGCGAGTGTATCACTCAATGTTATGAGGACATGCAATTAATCTCTACTGTCAAGCC 3195
Db      3360 GGGCGAGTGTATCACTCAATGTTATGAGGACATGCAATTAATCTCTACTGTCAAGCC 3419
Qy      3196 ATCCGTAAGATGCTTTCTGTGACTGTGAGTACTCAACCAAGTCAATCTGAGATAGTG 3255
Db      3420 ATCCGTAAGATGCTTTCTGTGACTGTGAGTACTCAACCAAGTCAATCTGAGATAGTG 3479
Qy      3356 TATCGCGGACCGAGTGTCTTGTCCCGGCTCAATACGGATTAATACCGGCGACATAG 3315
Db      3480 TATCGCGGACCGAGTGTCTTGTCCCGGCTCAATACGGATTAATACCGGCGACATAG 3539
Qy      3316 CAGAACTTTAAAGTGCATCATTTGAAAAAGTCTTCCGGGGGAAAACTCAAGAGAT 3375
Db      3540 CAGAACTTTAAAGTGCATCATTTGAAAAAGTCTTCCGGGGGAAAACTCAAGAGAT 3599
Qy      3376 CTTACCGCTGTGAGATCCAGTTGATGATTAACCACTCTGTGCAACCACTGATCTTCAAC 3435
Db      3600 CTTACCGCTGTGAGATCCAGTTGATGATTAACCACTCTGTGCAACCACTGATCTTCAAC 3659
Qy      3436 ATCTTTTACTTTCACAGCGTTTCTGCTGAGGAGCAAAAAACAGAAAGGCAAAATGCGCAA 3495
Db      3660 ATCTTTTACTTTCACAGCGTTTCTGCTGAGGAGCAAAAAACAGAAAGGCAAAATGCGCAA 3719
Qy      3496 AAAGGAAATTAAGGCGGACAGGAAATGTTGAATCTCATCTCTCTTTTTCAAATTA 3555
Db      3720 AAAGGAAATTAAGGCGGACAGGAAATGTTGAATCTCATCTCTCTTTTTCAAATTA 3779
Qy      3556 TTGAAGATTTATCAGGATTAATGTCTCATGAGCGGATACATATTTGAATTAATTAAGAA 3615
Db      3780 TTGAAGATTTATCAGGATTAATGTCTCATGAGCGGATACATATTTGAATTAATTAAGAA 3839
Qy      3616 AAATTAACAAATAGGGGTTCCGCGCACATTTCCCGAAAAAGTGCAC 3662
Db      3840 AAATTAACAAATAGGGGTTCCGCGCACATTTCCCGAAAAAGTGCAC 3886

```

Search completed: January 26, 2005, 08:25:58  
 Job time : 15663.5 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 26, 2005, 02:11:40 : Search time 11509 Seconds

(without alignments)  
11594.581 Million cell updates/sec

Title: US-10-826-523-41

Perfect score: 3662  
Sequence: 1 ctaaatgtgaagcgttaata.....attccccgaagaatgcccac 3662

Scoring table: IDENTITY\_NUC

Gapop 10.0, Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0\*

Listing first 45 summaries

Database :

EST:\*  
1: gb\_est1.\*  
2: gb\_est2.\*  
3: gb\_est3.\*  
4: gb\_est4.\*  
5: gb\_est5.\*  
6: gb\_est6.\*  
7: gb\_est6.\*  
8: gb\_g881.\*  
9: gb\_g882.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	1055.8	28.8	1070	1	AJ281552 4A3A-PeF1
2	986.4	26.9	1013	4	BM438846 Iplvtr0015
3	955.2	26.1	1067	1	AJ081137 AU081137
4	924	25.2	1004	1	AJ281480 4A3A-PaG8
5	923.8	25.2	1089	1	AU081124 AU081124
6	902.8	24.7	1049	9	CU021189 CH216-8A1
7	889.8	24.3	928	7	CO487414 G00227.B7
8	881.4	24.1	925	6	CB886151 Bn01b_020
9	870.2	23.8	935	4	BG838279 Gc01_10e0
10	865	23.6	1073	6	CF269652 FcY1Colid8
11	857.6	23.0	885	1	CU076016 CH216-138
12	841	23.0	841	1	AL042026 DXF2p434E
13	824.6	22.5	854	4	BM438850 Iplvtr0049
14	824	22.5	1169	9	AG332951 Mus muscu
15	821.4	22.4	1025	9	CU021193 CH216-8A1
16	820.2	22.4	865	7	CK125894 BES182411
17	819.6	22.4	1249	8	B2572284 msh2_2572
18	814.4	22.2	1126	8	B2577702 msh2_5533
19	809.4	22.1	856	7	CN823189 Oa_splbn
20	808	22.1	819	6	CD645375 Cvdgdn008
21	807.6	22.1	966	8	B2570738 msh2_1513
22	797.6	21.8	1011	8	B2576726 msh2_5071
23	793	21.7	827	7	CN823902 Oa_splbn
24	788.6	21.5	833	6	CB686421 Bn01b_04T

25	786.8	21.5	1163	1	AU081044 AU081044
26	785.4	21.4	1574	8	B2572566 msh2_2693
27	777.8	21.2	800	1	AJ281449 4A3A-PaD5
28	774.8	21.2	1336	6	B2575810 msh2_4637
29	772.4	21.1	789	6	CD280920 CD280920
30	772.4	21.1	954	1	AL044364 DXF2p434C
31	771.4	21.1	832	7	CN822433 Oa_splbn
32	767.6	21.0	863	7	CF752100 TCDB9_Hum
33	767.6	21.0	1370	8	B2571721 msh2_2025
34	763.2	20.8	780	5	BQ825693 10J0129B0
35	757	20.7	759	6	CD279661 CD279661
36	755.4	20.6	1463	8	B2571475 msh2_1906
37	752.4	20.5	1369	8	B2579291 msh2_6285
38	751.4	20.5	786	7	CN823164 Oa_splbn
39	750.4	20.5	752	6	CD646006 S8F6526R
40	746.4	20.4	1194	8	B2578985 msh2_6095
41	745.4	20.4	844	8	B2574513 msh2_3706
42	744.6	20.3	778	5	BP539067 BP539067
43	743.6	20.3	832	4	BG923768 602825893
44	741.4	20.2	758	9	CL422788 AE0544_Sa
45	741	20.2	741	6	CD279174 CD279174

## ALIGNMENTS

RESULT 1  
AJ281552 1070 bp mRNA linear EST 30-JUN-2000  
LOCUS 4A3A-PeF1-F Anopheles gambiae immune competent 4A3A Anopheles  
DEFINITION gambiae cDNA clone 4A3A-PeF1, mRNA sequence.  
ACCESSION AJ281552  
VERSION AJ281552.1 GI:6929432  
KEYWORDS EST.  
SOURCE Anopheles gambiae (African malaria mosquito)  
ORGANISM Anopheles gambiae  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae;  
Anopheles.  
REFERENCE 1 (bases 1 to 1070)  
Dimopoulos, G., Casavant, T.L., Chang, S., Scheetz, T., Robert, C.,  
Donohue, M., Schultz, J., Bense, V., Bork, P., Anorge, W., Soares, M.B.  
and Kafatos, F.C.  
Anopheles gambiae pilot gene discovery project: identification of  
mosquito innate immunity genes from expressed sequence tags  
generated from immune-competent cell lines  
Proc. Natl. Acad. Sci. U.S.A. 97 (12), 6619-6624 (2000)  
MEDLINE 20300950  
PUBMED 10841561  
COMMENT Contact: Dimopoulos G  
Fotis C. Kafatos Laboratory  
European Molecular Biology Laboratory  
Meyerhofstrasse 1, 69117 Heidelberg, Germany.  
TITLE Anopheles gambiae immune competent 4A3A  
JOURNAL Location/Qualifiers  
MEDLINE 1. 1070  
PUBMED /organism="Anopheles gambiae"  
/mol\_type="mRNA"  
/strain="4A r/r"  
/db\_xref="taxon:7165"  
/clone="4A3A-PeF1"  
/cell\_line="immune competent 4A3A"  
/lab\_host="E. coli DH10B"  
/clone\_lib="Anopheles gambiae immune competent 4A3A"  
/note="Vector: pT73D-Pac (Pharmacia) with a modified  
polylinker; Site 1: EcoRI; Site 2: NotI; sequenced from  
forward priming site which reads from the 3' end of the  
cDNA. The 4A3A is a directionally cloned and normalized  
cDNA library that was constructed from the 4A3A cell line  
oligo-T primed cDNA according to: Bonaldo, Lennon & Soares  
(1996) : Normalization and Subtraction: Two approaches to  
Facilitate Gene Discovery, Genome Research 6, 791-806."

ORIGIN



Query Match 28.8%; Score 1055.8; DB 1; Length 1070;  
 Best Local Similarity 99.7%; Pred. No. 3.8e-299;  
 Matches 1068; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

```

QY 1904 TTGCTGCGGTTTTCATAGAGCTCCGCCCTGAGAGATACAAAAAATGACCTCA 1963
DB 1 TTGCTGCGGTTTTCATAGAGCTCCGCCCTGAGAGATACAAAAAATGAGCCTCA 60
QY 1964 AGTCAGAGGTGGGAAACCCGACAGACTATTAAGATACAGGCGTTTCCCTGGAGC 2023
DB 61 AGTCAGAGGTGGGAAACCCGACAGACTATTAAGATACAGGCGTTTCCCTGGAGC 120
QY 2024 TCCCTGCGGCTCTCTGTTCCGACCTGCGCTTACCGGATACCTGTCGCTTCTC 2083
DB 121 TCCCTGCGGCTCTCTGTTCCGACCTGCGCTTACCGGATACCTGTCGCTTCTC 180
QY 2084 CCTTCGGGAAGCTGGCGCTTCTCATAGCTCAGCTGAGTATCTCAGTTCCGCTG 2143
DB 181 CCTTCGGGAAGCTGGCGCTTCTCATAGCTCAGCTGAGTATCTCAGTTCCGCTG 240
QY 2144 GTGCTTCGCTCCAGCTGGGCTGTGTGCAAGAACCCCGTTCAAGCCGACGCTGCGC 2203
DB 241 GTGCTTCGCTCCAGCTGGGCTGTGTGCAAGAACCCCGTTCAAGCCGACGCTGCGC 300
QY 2204 TTATCCGCTACTATCTCTTGAATCCAAACCGGATACAGACTTATCCCACTGCA 2263
DB 301 TTATCCGCTACTATCTCTTGAATCCAAACCGGATACAGACTTATCCCACTGCA 360
QY 2264 GCAGCCACTGTAAACAGATTAAGACAGAGATTAAGGCGGTGCTACAGAGTTTGT 2323
DB 361 GCAGCCACTGTAAACAGATTAAGACAGAGATTAAGGCGGTGCTACAGAGTTTGT 420
QY 2324 AAGTGTGGCTTAACCTACGCGCTACACTAAGAGACAGATTTGCTATCTGCGCTG 2383
DB 421 AAGTGTGGCTTAACCTACGCGCTACACTAAGAGACAGATTTGCTATCTGCGCTG 480
QY 2384 AAGCAGTTACCTTCGGAAGAAAGTGTGCTGCTTGAATCCGCAAAACACACGCT 2443
DB 481 AAGCAGTTACCTTCGGAAGAAAGTGTGCTGCTTGAATCCGCAAAACACACGCT 540
QY 2444 GGTAGCGGTGTTTTTTTGTTCAGACAGAGATTAAGGCGAGAAAGATTCGA 2503
DB 541 GGTAGCGGTGTTTTTTTGTTCAGACAGAGATTAAGGCGAGAAAGATTCGA 600
QY 2504 GAAGATCCTTGAATCTTTCAGCGGCTGAGGCTCAGTGAAGCAAACTCAGCTTAA 2563
DB 601 GAAGATCCTTGAATCTTTCAGCGGCTGAGGCTCAGTGAAGCAAACTCAGCTTAA 660
QY 2564 GGGATTTTGTATGATTAATCAAAAAGATCTTCACCTAGATCTTTTAAATTA 2623
DB 661 GGGATTTTGTATGATTAATCAAAAAGATCTTCACCTAGATCTTTTAAATTA 720
QY 2624 TGAAGTTTAAATCAATCTAAAGTATATAGTAAACTTGTCTGACAGTTACCAATGC 2683
DB 721 TGAAGTTTAAATCAATCTAAAGTATATAGTAAACTTGTCTGACAGTTACCAATGC 780
QY 2684 TTAATCAGTAGGACCACTATCTCAGGATCTGTCTATTCGTCATCTCAATAGTCC 2743
DB 781 TTAATCAGTAGGACCACTATCTCAGGATCTGTCTATTCGTCATCTCAATAGTCC 840
QY 2744 CTCCTCGGTGTATGATTAATCAATGAGGAGGCTTACATCTGAGGCGCAGTCTGCA 2803
DB 841 CTCCTCGGTGTATGATTAATCAATGAGGAGGCTTACATCTGAGGCGCAGTCTGCA 900
QY 2804 ATGATATCCGAGAACCCGCTCACCGGCTCCAGTTTATCAGCAATTAACCAAGCACC 2863
DB 901 ATGATATCCGAGAACCCGCTCACCGGCTCCAGTTTATCAGCAATTAACCAAGCACC 960
QY 2864 GGAAGGCGCAGAGCAGAAAGTGTCTGCAACTTATCCGCTCAGTCAATTAAT 2923
DB 961 GGAAGGCGCAGAGCAGAAAGTGTCTGCAACTTATCCGCTCAGTCAATTAAT 1019
QY 2924 TGTTCGCGGAAGCTAAGATAGTTCGCGCAATTAATGTTGGCAAC 2974
  
```

DB 1020 TGTTCGCGGAAGCTAAGATAGTTCGCGCAATTAATGTTGGCAAC 1070

RESULT 2  
 BM438846/c 1013 bp mRNA linear EST 31-JAN-2002  
 LOCUS Iplv00157 Liver cDNA library Ictalurus punctatus cDNA 5', mRNA  
 DEFINITION  
 BM438846  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Ictalurus punctatus (channel catfish)  
 Ictalurus punctatus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Siluriformes;  
 Ictaluridae; Ictalurus.  
 1 (bases 1 to 1013)  
 REFERENCE  
 Feng, J., Kucuktas, H., Kocabas, A., Li, P. and Liu, Z.  
 Transcription of channel catfish (Ictalurus punctatus): initial  
 analysis of expressed sequence tags from the liver  
 JOURNAL  
 Unpublished (2002)  
 COMMENT  
 Contact: Liu ZJ  
 The Fish Molecular Genetics and Biotechnology Laboratory,  
 Department of Fisheries and Allied Aquacultures and Program of Cell  
 and Molecular Biosciences  
 Auburn University  
 203 Swingle Hall, Auburn University, Auburn, AL 36849, USA  
 Tel: 334 844 4054  
 Fax: 334 844 9208  
 Email: zliu@acesag.auburn.edu  
 Seg primer: M13 Reverse.  
 Location/Qualifiers  
 1..1013  
 /organism="Ictalurus punctatus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:7998"  
 /clone\_id="Liver cDNA library"  
 /note="Organ: Liver; Vector: pSPORT1; Site\_1: NotI;  
 Site\_2: SalI"

#### FEATURES

source  
 1..1013  
 /organism="Ictalurus punctatus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:7998"  
 /clone\_id="Liver cDNA library"  
 /note="Organ: Liver; Vector: pSPORT1; Site\_1: NotI;  
 Site\_2: SalI"

Query Match 26.9%; Score 986.4; DB 4; Length 1013;  
 Best Local Similarity 99.5%; Pred. No. 1.1e-278;  
 Matches 1009; Conservative 1; Mismatches 2; Indels 2; Gaps 2;

```

QY 2477 ATTACGGCAGAAAAAGGATCTCAAGAGATCTTGAATCTTTCTACGGGCTGTAC 2536
DB 1013 ATTACGGCAGAAAAAGGATCTCAAGAGATCTTGAATCTTTCTACGGGCTGTAC 954
QY 2537 GCTCAGTGAACGAAAACTCAGTTAAGGATTTGGTCATGAGATTATCAAAAGATC 2596
DB 953 GCTCAGTGAACGAAAACTCAGTTAAGGATTTGGTCATGAGATTATCAAAAGATC 894
QY 2597 TTCACCTAGATCTTTTAAATTAAGATTTTAAATCAATCTAAAGTATATATAG 2656
DB 893 TTCACCTAGATCTTTTAAATTAAGATTTTAAATCAATCTAAAGTATATATAG 834
QY 2657 TAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAAGCACTATCTCAGCGAT 2716
DB 833 TAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAAGCACTATCTCAGCGAT 774
QY 2717 CTATTTGTTCAATCCATGTTGCTCGATCCCGCTGCTGTAGTATCACTACATACG 2776
DB 773 CTATTTGTTCAATCCATGTTGCTCGATCCCGCTGCTGTAGTATCACTACATACG 714
QY 2777 GGTTCACATCTTGCCCAAGTGTGCAATGATTAACGACCAACGCTCACCGGCTCCA 2836
DB 713 GGTTCACATCTTGCCCAAGTGTGCAATGATTAACGACCAACGCTCACCGGCTCCA 654
QY 2837 GATTTATCAGCAATTAACACGACCGGCAAGGCGCAGACGCAAGTGTCTGCAACT 2896
DB 653 GATTTATCAGCAATTAACACGACCGGCAAGGCGCAGACGCAAGTGTCTGCAACT 594
  
```



2897 TTATCCGCTCCATCCAGTCTATTAAATTGTGCGGGAAGCTAGAGTAAGTTCGCCA 2956  
Db TTATCCGCTCCATCCAGTCTATTAAATTGTGCGGGAAGCTAGAGTAAGTTCGCCA 534  
2957 GTTAATAGTTCGCAAGTTCGTCATGCTACAGGACATGCTGTCACGCTCGCG 3016  
Db GTTAATAGTTCGCAAGTTCGTCATGCTACAGGACATGCTGTCACGCTCGCG 474  
3017 TTTGGATAGCTTCACTACGCTCCGCTTCCCAAGATCAAGCGGAGTTACA-TGATCCCG 3075  
Db TTTGGATAGCTTCACTACGCTCCGCTTCCCAAGATCAAGCGGAGTTACA-TGATCCCG 414  
3076 CATGTGCGCAAAAAGCGGTAGCTCTGCTGCTCCGATCCGATCCGTCGCAAGAAAGTT 3135  
Db CATGTGCGCAAAAAGCGGTAGCT-CTTCGCTCTCCGATCCGTCGCAAGAAAGTT 355  
3136 GGCCGAGTGTATCACTACATGATGATGAGCAGACTGATATTTCTTTACTGTCATGCC 3195  
Db GGCCGAGTGTATCACTACATGATGATGAGCAGACTGATATTTCTTTACTGTCATGCC 295  
3196 ATCCGTAAGATGCTTTTCTGTGACTGTGAGTACTCAACCAAGTCATTCTGAGAATAGTG 3255  
Db ATCCGTAAGATGCTTTTCTGTGACTGTGAGTACTCAACCAAGTCATTCTGAGAATAGTG 235  
3256 TATGCGGAGACCGAGTGTCTCTTCCGCGCTCAATACGGATTAATACCGGCGCAATAG 3315  
Db TATGCGGAGACCGAGTGTCTCTTCCGCGCTCAATACGGATTAATACCGGCGCAATAG 175  
3316 CAGAACTTTAAAGATGCTCATCATTTGCAAAAGTTCTTCCGCGGCAAACTCTCAAGAT 3375  
Db CAGAACTTTAAAGATGCTCATCATTTGCAAAAGTTCTTCCGCGGCAAACTCTCAAGAT 115  
3376 CTTACCGCTGTGAGATCAGTTCGATTAACCACTGTCGACCCAACTGATCTTACG 3435  
Db CTTACCGCTGTGAGATCAGTTCGATTAACCACTGTCGACCCAACTGATCTTACG 55  
3436 ATCTTTAACTTTACCGAGCTTCTGCTGAGCAAAAACAGAAAGGCAAAATGC 3489  
Db ATCTTTAACTTTACCGAGCTTCTGCTGAGCAAAAACAGAAAGGCAAAATGC 1

RESULT 3  
A081137 1067 bp mRNA linear EST 30-JUL-2002  
LOCUS A081137 Oncorhynchus mykiss Kidney infected by infectious  
DEFINITION hematopoietic necrosis virus Oncorhynchus mykiss cDNA clone K12,  
mRNA sequence.  
ACCESSION A081137  
VERSION A081137.1 GI:6431485  
KEYWORDS EST.  
SOURCE Oncorhynchus mykiss (rainbow trout)  
ORGANISM Oncorhynchus mykiss  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Actinopterygii; Neopterygii; Teleostei; Euteleostei;  
Proteanchopeterygii; Salmoniformes; Salmonidae; Oncorhynchus.  
1 (bases 1 to 1067)  
Kono, T., Sakai, M. and Lapetra, S.B.  
Expressed Sequence Tag Analysis of Kidney and Gill Tissues from  
Rainbow Trout (Oncorhynchus mykiss) Infected with Infectious  
Hematopoietic Necrosis Virus  
Mar. Biotechnol. 2 (5), 493-498 (2001)  
Contact: Masahiro Sakai  
Faculty of Agriculture  
Miyazaki University  
1-1 nishi gakuendibandai, Miyazaki, Miyazaki 889-2192, Japan  
Email: m.sakai@cc.miyazaki-u.ac.jp.  
location/Qualifiers  
1..1067  
/organism="Oncorhynchus mykiss"  
/mol\_type="mRNA"  
/db\_xref="taxon:8022"  
/clone="K12"

ORIGIN  
/issue\_type="kidney"  
/clone\_id="Oncorhynchus mykiss kidney infected by  
infectious hematopoietic necrosis virus"  
/note="common name:rainbow trout ; infected by infectious  
hematopoietic necrosis virus"  
Query Match 26.1%; Score 955.2; DB 1; Length 1067;  
Best Local Similarity 99.0%; Pred. No. 1.7e-269;  
Matches 982; Conservative 0; Mismatches 8; Indels 2; Gaps 2;  
1494 CGCCCTTGCGCTAATCATGATGATGCTGTTCTCTGATGAATTTATCCGTCACAA 1553  
Db CGACCTTGCGCTAATCATGATGATGCTGTTCTCTGATGAATTTATCCGTCACAA 137  
1554 TTCCACACAAATAGAGCGGAGCATTAAGTAAAGCTGGGAGTCCATAGATGA 1613  
Db TTCCACACAAATAGAGCGGAGCATTAAGTAAAGCTGGGAGTCCATAGATGA 197  
1614 GCTAATCACTTAATTTGCTTGGCTCACTGCCCCCTTTCAGTCCGGAACCTGTGCT 1673  
Db GCTAATCACTTAATTTGCTTGGCTCACTGCCCCCTTTCAGTCCGGAACCTGTGCT 257  
1674 GCCAGCTGCATTAATGAATCGGCCAAGCGCGGGAAGAGCGGTTTGATTTGGCGCT 1733  
Db GCCAGCTGCATTAATGAATCGGCCAAGCGCGGGAAGAGCGGTTTGATTTGGCGCT 317  
1734 CTTCCGCTCTCTGCTCACTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1793  
Db CTTCCGCTCTCTGCTCACTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 377  
1794 CAGCTCACTCAAAAGCGGTAAATAGGTTATTCACAGAAATCAGGGGATTAACAGAAAGA 1853  
Db CAGCTCACTCAAAAGCGGTAAATAGGTTATTCACAGAAATCAGGGGATTAACAGAAAGA 437  
1854 ACATGTAGCAAAAGCGCAGCAAAAGCGCAGAAACGTAAGAAAGCGGCTGTCGCGCT 1913  
Db ACATGTAGCAAAAGCGCAGCAAAAGCGCAGAAACGTAAGAAAGCGGCTGTCGCGCT 497  
1914 TTTTCATAGGCTCCGCCCTCTGACGAGATCAAAATTCGACGCTCAAGTCAAGT 1973  
Db TTTTCATAGGCTCCGCCCTCTGACGAGATCAAAATTCGACGCTCAAGTCAAGT 557  
1974 GCGCAAAACCGACAGGACTTAAGATACAGAGCGTTTCCCTCGAAAGCTCCTCGTGC 2033  
Db GCGCAAAACCGACAGGACTTAAGATACAGAGCGTTTCCCTCGAAAGCTCCTCGTGC 617  
2034 GCTCTCGTTCGCAACCTGCGCTTACCGGATACCTGTGCGGCTTCTCCCTCGGAA 2093  
Db GCTCTCGTTCGCAACCTGCGCTTACCGGATACCTGTGCGGCTTCTCCCTCGGAA 677  
2094 GCGTGGCGCTTTCATAGTCAAGCTGATGATATTCAGTTCGATGAGTTCGCTGCT 2153  
Db GCGTGGCGCTTTCATAGTCAAGCTGATGATATTCAGTTCGATGAGTTCGCTGCT 737  
2154 CCAAGCTGGCTGTGTGCAAGAACCCCGCTTACGCCGACCGCTGAGCTTATCCGCTA 2213  
Db CCAAGCTGGCTGTGTGCAAGAACCCCGCTTACGCCGACCGCTGAGCTTATCCGCTA 797  
2214 ACTATGCTTTGAGTCAACCCCGGTAGACACGCTTATGCGCACTGCGACGACCACTG 2273  
Db ACTATGCTTTGAGTCAACCCCGGTAGACACGCTTATGCGCACTGCGACGACCACTG 857  
2274 GTAAACAGATTAGCAGAGCAGATGATGAGCGGTGCTACAGAGTTCTTGAAGTGTGCG 2333  
Db GTAAACAGATTAGCAGAGCAGATGATGAGCGGTGCTACAGAGTTCTTGAAGTGTGCG 917  
2334 CTTAATCGGCTACATGAGAGACAGATTTTGTATCTGCGCTCTGTAAGCACTTA 2393  
Db CTTAATCGGCTACATGAGAGACAGATTTTGTATCTGCGCTCTGTAAGCACTTA 977  
2394 CCTTCGGAAGAGGTTGATGCTTATGCTTACCGCAACCAACGCTGTGCGGT 2453  
Db CCTTCGGAAGAGGTTGATGCTTATGCTTACCGCAACCAACGCTGTGCGGT

Db 978 CTTGCGAAGAGTTGAGTCTTGATCCGGCAACAA-CAGCGTGAGCGTGT 1036

QY 2454 GTTTTGTGTCAGACGAGATTACGCC 2485

Db 1037 TTTTGTGTTGC-AGCAGCAGATTACGCC 1067

RESULT 4

AJ281480 1004 bp mRNA linear EST 30-JUN-2000

LOCUS 4A3A-P4G8-F Anopheles gambiae immune competent 4A3A Anopheles

DEFINITION gambiae CDNA clone 4A3A-P4G8, mRNA sequence.

ACCESSION AJ281480

VERSION AJ281480.1 GI:6929360

KEYWORDS EST.

SOURCE Anopheles gambiae (African malaria mosquito)

ORGANISM Anopheles gambiae

REFERENCE Neopletia; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Eukaryota; Endopterygota; Diptera; Nematocera; Culicoidae; Anopheles.

AUTHORS Dimopoulos, G., Caeavant, T.L., Chang, S., Schaefer, T., Roberts, C., Donohue, M., Schultz, J., Benes, V., Bork, P., Ansoerge, W., Soares, M.B. and Kafatos, F.C.

TITLE Anopheles gambiae pilot gene discovery project: identification of mosquito innate immunity genes from expressed sequence tags

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (12), 6619-6624 (2000)

MEDLINE 20300950

PUBMED 10841561

COMMENT Contact: Dimopoulos G  
Fotis C. Kafatos Laboratory  
European Molecular Biology Laboratory  
Meyerhofstrasse 1, 69117 Heidelberg, Germany.  
Location/Qualifiers  
1..1004  
/organism="Anopheles gambiae"  
/mol\_type="mRNA"  
/strain="4A r/r"  
/db\_xref="taxon:7165"  
/clone="4A3A-P4G8"  
/cell\_line="Immune competent 4A3A"  
/lab\_host="E. coli DH10B"  
/note="Vector: pUT73D-pac (Pharmacia) with a modified polylinker; Site\_1: EcoRI; Site\_2: NotI; sequenced from forward priming site which reads from the 3' end of the CDNA. The 4A3A is a directionally cloned and normalized CDNA library that was constructed from the 4A3A cell line (1996) : Normalization and Subtraction: Two approaches To Facilitate Gene Discovery, Genome Research 6, 791-806."

ORIGIN

Query Match 25.2%; Score 924; DB 1; Length 1004;  
Best Local Similarity 99.0%; Pred. No. 2.7e-260;  
Matches 1001; Conservative 1; Mismatches 2; Indels 7; Gaps 7;

QY 1875 AAAAGCCAGGAACCGTAAAGAGCGGCTGCTGGCTTTTCCATNAGCTCCGCC 1934

Db 1 AAAAGCCAGGAACCGTAAAGAGCGGCTGCTGGCTTTTCCATNAGCTCCGCC 60

QY 1935 CTGACGAGCATCAAAAATCGACGCTCAAGTCAGAGGTGGGAAACCGCAGAGACTAT 1994

Db 61 CTGACGAGCATCAAAAATCGACGCTCAAGTCAGAGGTGGGAAACCGCAGAGACTAT 120

QY 1995 AAAAGTACAGGCGTTTCCCGTGAAGCTCCCTGTGCGCTTCTCTGTTCCGACCTGC 2054

Db 121 AAAAGTACAGGCGTTTCCCGTGAAGCTCCCTGTGCGCTTCTCTGTTCCGACCTGC 180

QY 2055 CGCTTACCGGATACCTGTGCGCTTCTCTGCGGAAAGGTGGCGCTTCTCATNAGCT 2114

Db 181 CGCTTACCGGATACCTGTGCGCTTCTCTGCGGAAAGGTGGCGCTTCTCATNAGCT 240

QY 2115 CACGCTGAGTATCTCAGTTCGGTGTAGGTGCTGCTCCAGAGCTGGGCTGTGACAG 2174

Db 241 CACGCTGAGTATCTCAGTTCGGTGTAGGTGCTGCTCCAGAGCTGGGCTGTGACAG 300

QY 2175 AACCCCGCTTCAAGCCGACCGCTGCGCTTATCCGTAATCTATCGTCTTGAAGTCCAAC 2234

Db 301 AACCCCGCTTCAAGCCGACCGCTGCGCTTATCCGTAATCTATCGTCTTGAAGTCCAAC 360

QY 2235 CGGTAAAGACAGACTTATCGCCACCTGGAGCGGCACTGGTAACAGATTTAGCAGACGA 2294

Db 361 CGGTAAAGACAGACTTATCGCCACCTGGAGCGGCACTGGTAACAGATTTAGCAGACGA 420

QY 2295 GGTATGTAGGCGGTGCTCAAGAGTCTTGAAGTGTGGCTTAATCAAGCTTACACTAGAA 2354

Db 421 GGTATGTAGGCGGTGCTCAAGAGTCTTGAAGTGTGGCTTAACTAAGCTTACACTAGAA 480

QY 2355 GGACAGTATTTGGTATCTGCGCTGCTGAGACCCGTTACTTGGAAAAAGTTGGTA 2414

Db 481 GGACAGTATTTGGTATCTGCGCTGCTGAGACCCGTTACTTGGAAAAAGTTGGTA 540

QY 2415 GCTCTTGATCGGCAACAAACACCGCTGTAGCGGTGTTTGTGTTGCAAGACG 2474

Db 541 GCTCTTGATCGGCAACAAACACCGCTGTAGCGGTGTTTGTGTTGCAAGACG 600

QY 2475 AGATTACGGCGGAAAAAGGATCTCAAGAGTCTTGTATCTTTTCTACGGGCTGT 2534

Db 601 AGATTACGGCGGAAAAAGGATCTCAAGAGTCTTGTATCTTTTCTACGGGCTGT 660

QY 2535 ACCTCAGTGAACGAAACCTACGTTAAGGATTTTGTATGATTAATCAAAAGGA 2594

Db 661 ACCTCAGTGAACGAAACCTACGTTAAGGATTTTGTATGATTAATCAAAAGGA 720

QY 2595 TCTTCACTTATGCTTTTAAATTAAGATTTTAAATCAATCAATCAATCAATCAATCAAT 2654

Db 721 TCTTCACTTATGCTTTTAAATTAAGATTTTAAATCAATCAATCAATCAATCAATCAAT 780

QY 2655 AGTAACTGTGTGACAGTATCAATGCTTAATCAATGAGGACCTATCTGAGCATCT 2714

Db 781 AGTAACTGTGTGACAGTATCAATGCTTAATCAATGAGGACCTATCTGAGCATCT 839

QY 2715 GTCTATTGCTTATCCATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2774

Db 840 GTCTATTGCTTATCCATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 898

QY 2775 AGGCTTACCATGTCGCCAGTGTGCAATGATTAACGGAGACCAAGCTACCGGCTC 2834

Db 899 AGGCTTACCATGTCGCCAGTGTGCAATGATTAACGGAGACCAAGCTACCGGCTC 955

QY 2835 CAGATTATCAGAAATTAACAGCCAGCCGGAAGGCGCAGGCGAAGTG 2885

Db 956 CAGATTATCAGAAATTAACAGCCAGCCGGAAGGCGCAGGCGAAGTG 1004

RESULT 5

AU081124 1089 bp mRNA linear EST 30-JUL-2002

LOCUS AU081124

DEFINITION AU081124 Oncorhynchus mykiss Kidney infected by infectious hematopoietic necrosis virus Oncoirhynchus mykiss CDNA clone Kc12, mRNA sequence.

ACCESSION AU081124.1 GI:6431472

VERSION AU081124

KEYWORDS EST.

SOURCE Oncoirhynchus mykiss (rainbow trout)

ORGANISM Oncoirhynchus mykiss

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Protacanthopterygii; Salmoniformes; Salmonidae; Oncoirhynchus. 1 (bases 1 to 1089)

AUTHORS Kono, T., Sakai, M. and LaPatra, S.E.

TITLE Expressed Sequence Tag Analysis of Kidney and Gill Tissues from Rainbow Trout (Oncoirhynchus mykiss) Infected with Infectious Hematopoietic Necrosis Virus

JOURNAL Mar.: Biotechnol. 2 (5), 493-498 (2001)  
Contact: Masahito Sakai

## COMMENT

Faculty of Agriculture  
Miyazaki University  
1-1 nishi gakuenkibanadai, Miyazaki, Miyazaki 889-2192, Japan  
Email: m.sakai@cc.miyazaki-u.ac.jp.  
Location/Qualifiers

## FEATURES

source

1. 1089  
/organism="Oncoerhynchus mykiss"  
/mol\_type="mRNA"  
/db\_xref="taxon:8022"  
/clone="KC12"  
/issue\_type="kidney"  
/clone\_lib="Oncoerhynchus mykiss kidney infected by  
infectious hematopoietic necrosis virus"  
/note="common name:rainbow trout ; infected by infectious  
hematopoietic necrosis virus"

## ORIGIN

Query Match 25.2%; Score 923.8; DB 1; Length 1089;

Best Local Similarity 97.2%; Pred. No. 3.2e-260;  
Matches 983; Conservative 0; Mismatches 22; Indels 6; Gaps 4;

QY 1494 CGCCCTGGCGTAAATCATAGTCATAGCTGTTCTGTGAAATGTTATCCGCTCACAA 1553  
DB 79 CGAGCTTGCGCTAATCATAGTCATAGCTGTTCTGTGAAATGTTATCCGCTCACAA 138  
QY 1554 TTCCACACATACGAGCCGGAAGCATAAAGTAAAGCTGGGGTCCCTAATGATGA 1613  
DB 139 TTCCACACATACGAGCCGGAAGCATAAAGTAAAGCTGGGGTCCCTAATGATGA 198  
QY 1614 GCTAATCACAATTAATGCGTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1673  
DB 199 GCTAATCACAATTAATGCGTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 258  
QY 1674 GCCAGCTGATTAATGAAATCGGCCCAACGCGGGGAGAGGCGTTTGGTATTGGGCGCT 1733  
DB 259 GCCAGCTGATTAATGAAATCGGCCCAACGCGGGGAGAGGCGTTTGGTATTGGGCGCT 318  
QY 1734 CTTCCGCTCTCTCGCTCACTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1793  
DB 319 CTTCCGCTCTCTCGCTCACTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 378  
QY 1794 CAGCTCACTCAAAAGCGGTATAGGTTATCCACAGAAATCGAGGATTAACGAGAAAGA 1853  
DB 379 CAGCTCACTCAAAAGCGGTATAGGTTATCCACAGAAATCGAGGATTAACGAGAAAGA 438  
QY 1854 ACATGTGAGCAAAAGGCGAGCAAAAGGCGAGAAACCGTAAAGGCGGCTGCGGCGCT 1913  
DB 439 ACATGTGAGCAAAAGGCGAGCAAAAGGCGAGAAACCGTAAAGGCGGCTGCGGCGCT 498  
QY 1914 TTTTCATAGGCTCGCGCCCTGAGCGAGATCAAAAAATCGAGCTCAAGTCAGAGT 1973  
DB 499 TTTTCATAGGCTCGCGCCCTGAGCGAGATCAAAAAATCGAGCTCAAGTCAGAGT 558  
QY 1974 GGCAGAAACCGGAGAGTAAAGATACAGAGGTTTCCCTCGAAGCTCCCTCGTGC 2033  
DB 559 GGCAGAAACCGGAGAGTAAAGATACAGAGGTTTCCCTCGAAGCTCCCTCGTGC 618  
QY 2034 GCTCTCGTGTCCGACCTGCGGCTTACCGGATACCTGCGGCTTTCCTTCGAGAA 2093  
DB 619 GCTCTCGTGTCCGACCTGCGGCTTACCGGATACCTGCGGCTTTCCTTCGAGAA 678  
QY 2094 GCGTGGCGCTTTCATAGCTCAAGCTGATGATCTCAAGTTCCGTTAGGTGCTGCT 2153  
DB 679 GCGTGGCGCTTTCATAGCTCAAGCTGATGATCTCAAGTTCCGTTAGGTGCTGCT 738  
QY 2154 CCAAGCTGGCTGTGTGACAGAACCCCGCTTACGCGGAGCCGCTGCGCTTATCCGTA 2213  
DB 739 CCAAGCTGGCTGTGTGACAGAACCCCGCTTACGCGGAGCCGCTGCGCTTATCCGTA 798  
QY 2214 ACTATCGTCTTGAATCCAAACCGGATGAAGACAGACTTATCGGCACTGAGCAGACGACTG 2273

DB 799 ACTATCGTCTTGAATCCAAACCGGATGAAGACAGACTTATCGGCACTGAGCAGACGACTG 858  
QY 2274 GTAACAGATTTAGCAGAGCGAGATGATGAGCGGTCTACAGAGTTCTTGAAG-TGCTGG 2332  
DB 859 GTAACAGATTTAGCAGAGCGAGATGATGAGCGGTCTACAGAGTTCTTGAAGTTTGGG 918  
QY 2333 CCTACTACGCGCTACATAGAGAGAGAGTATTTGTTATCTGCGCTGCTGAGCCAGTT 2392  
DB 919 CCTACTACGCGCTACATAGAGAGAGAGTATTTGTTATCTGCGCTGCTGAGCCAGTT 978  
QY 2393 ACCTTCGAAAAAGATGTTAGTCTT-GATCCGGCAACAAACA---CCGCTGTAG 2448  
DB 979 ACCTTCGAAAAAGATGTTAGTCTTGTGATTCGGCAACAAACAACCCCTGGAGAG 1038  
QY 2449 CCGTGTGTTTTTTTGTTCGAAGCA-GCAGATTACCGCGAGAAAAAAGAT 2498  
DB 1039 GGGGTGCTTTTGTGTTGCAAGCAGCAGATTAACCCGAAAAAAGGAT 1089

## RESULT 6

CL021189

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

CL021189 1049 bp DNA linear GSS 31-DEC-2003  
CH216-8A14 RM1.1 CH216 Xenopus tropicalis genomic clone CH216-8A14,  
genomic survey sequence.  
CL021189  
CL021189.1 GI:40463002  
GSS.  
Xenopus tropicalis (western clawed frog)  
Xenopus tropicalis  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;  
Xenopodinae; Xenopus; Silurana.  
Kremetzki, C., Carter, J., McPherson, J., Warren, W., Graves, T.,  
Mardis, E. and Wilson, R.  
A physical map of the xenopus tropicalis genome  
Unpublished (2003)  
Contact: Richard K Wilson  
Genome Sequencing Center  
Washington University School of Medicine  
Email: submissions@wustl.edu  
Insert length: 175000 Std Error: 0.00  
Seq primer: RM1.TACGACTCACTATAGGAGCA  
Class: BAC ends  
High quality sequence start: 43  
High quality sequence stop: 888.  
Location/Qualifiers

## FEATURES

source

1. 1049  
/organism="Xenopus tropicalis"  
/mol\_type="genomic DNA"  
/strain="Nigerian frog"  
/db\_xref="taxon:8364"  
/clone="CH216-8A14"  
/sex="male"  
/cell\_line="Stock 248 F7A2, inbred N7"  
/clone\_lib="CH216"  
/note="Vector: pTARBAC2.1; CHORI-216 Xenopus tropicalis  
BAC library"

## ORIGIN

Query Match 24.7%; Score 902.8; DB 9; Length 1049;  
Best Local Similarity 98.0%; Pred. No. 4.9e-254;  
Matches 935; Conservative 0; Mismatches 17; Indels 2; Gaps 2;  
QY 1486 CCGGCGGCGCGCTGCGTAAATCATAGTATGCTGTTCTGTGAAATGTTATCC 1545  
DB 95 CAGGATGCAAGCTGCGCTAATCATAGTATGCTGTTCTGTGAAATGTTATCC 154  
QY 1546 GCTCAATTCACACAAATAGAGCGGAGCATTAAGTAAAGCTGGGGTGCCTA 1605  
DB 155 GCTCAATTCACACAAATAGAGCGGAGCATTAAGTAAAGCTGGGGTGCCTA 214  
QY 1606 ATGAGTAGCTAATCATTAATGCGTTGCGCTACCTGCGGCTTTCAGATGCGGAAA 1665

	Db	215	ATGATGAGCTAACTCAATTAATTTGGTGGCGCTCACTGCCCCGCTTTCAGTCGGGAAA	274
QY	1666	CCTGTCTGTGCCAGCTGCATTAATGATCGGCAACGCGCGGGAGAGGGGTTTCGAT	1725	
Db	275	CCTGTCTGTGCCAGCTGCATTAATGATCGGCAACCGCGGGAGAGGGGTTTCGAT	334	
QY	1726	TGGGCGCTCTTCGCGCTTCTCTGCTCACTGACTGCTGCGCTGCTGCTGCTGCGCTG	1785	
Db	335	TGGGCGCTCTTCGCGCTTCTCTGCTCACTGACTGCTGCGCTGCTGCTGCTGCGCTG	394	
QY	1786	AGCGGTATCAGCTCACTCAAAAGCGGTAATACGGTTATCCAGAAATCAGGGGATAAGC	1845	
Db	395	AGCGGTATCAGCTCACTCAAAAGCGGTAATACGGTTATCCAGAAATCAGGGGATAAGC	454	
QY	1846	AGGAAAATAATGTGAGCAAAAAGCCAGCAAAAAGCCAGGAAACCTTAAAAAGCCGCTT	1905	
Db	455	AGGAAAATAATGTGAGCAAAAAGCCAGCAAAAAGCCAGGAAACCTTAAAAAGCCGCTT	514	
QY	1906	GCTGCGCTTTTTCATATAGGCTCCGCCCTCAGCAGCATCAAAAAATCAGCGCTCAG	1965	
Db	515	GCTGCGCTTTTTCATATAGGCTCCGCCCTCAGCAGCATCAAAAAATCAGCGCTCAG	574	
QY	1966	TCAAGGTGCGAAACCCGACAGGACTAATAAGATACAGGCGTTTCCCTTGAAGCTC	2025	
Db	575	TCAAGGTGCGAAACCCGACAGGACTAATAAGATACAGGCGTTTCCCTTGAAGCTC	634	
QY	2026	CCTCGTGCCTCTCTCTGTTCCGACCCTGCGCTTACCGATACCTGTCCGCTTTCGCC	2085	
Db	635	CCTCGTGCCTCTCTCTGTTCCGACCCTGCGCTTACCGATACCTGTCCGCTTTCGCC	694	
QY	2086	TTTCGGGAAACGTGCGGCTTTTCATATAGCTACGCGTATAGGATATTCAGTTTCGTTAGT	2145	
Db	695	TTTCGGGAAACGTGCGGCTTTTCATATAGCTACGCGTATAGGATATTCAGTTTCGTTAGT	754	
QY	2146	CGTTGCTCTCAAGCTGGGCTGTGTGCAGAAACCCCGGTTACAGCCGACGCTGCGCTT	2205	
Db	755	CGTTGCTCTCAAGCTGGGCTGTGTGCAGAAACCCCGGTTACAGCCGACGCTGCGCTT	814	
QY	2206	ATTCGGTAACTATCTCTTTAGTCCAAACCGGGTAAACAGACTTATCCGACCTGCGACG	2265	
Db	815	ATTCGGTAACTATCTCTTTAGTCCAAACCGGGTAAACAGACTTATCCGACCTGCGACG	874	
QY	2266	AGCCACTGTGTAAACAGGATTACAGACGAGGTATGAGGCGGTGCTACAGAGTTCTTGA	2325	
Db	875	AGCCACTGTGTAAACAGGATTACAGACGAGGTATGAGGCGGTGCTACAGAGTTCTTGA	2384	
QY	2326	GTGTGGGCTTAACAGGCTACGCTACGTAAAGGAC-AGATTTGGTATCTGCGCTGTCTGA	2384	
Db	935	GTGTGGGCTTAACAGGCTACGCTACGTAAAGGAC-AGATTTGGTATCTGCGCTGTCTGA	994	
QY	2385	AGCCAGTTAC-TTTCGAAAAAAGATTGTAGTCTTGTATCCGGCAAAACAAC 2437		
Db	995	AACCACTTACCTTTCGAAAAAAGATTGTAGTCTTGTATCCGGCAAAACAAC 1048		
RESULT 7				
LOCUS	CO487414	928 bp	mRNA	linear
DEFINITION	GM0227.B7.1 K04 GM022: ROOT XYLEM - mature trees Picea glauca cDNA			
ACCESSION	CO487414			
VERSION	CO487414.1	GI:50174316		
KEYWORDS				
SOURCE				
ORGANISM				
REFERENCE				
AUTHORS				
TITLE				

JOURNAL  
COMMENT

Unpublished (2004)  
Contact: John Mackay  
Centre de Recherche en Biologie Forestiere  
Universite Laval  
Pavillon Charles-Eugene Marchand, Quebec, Quebec, CANADA G1K 7P4  
Fax: 418 656 7493  
Email: jmackay@rsvs.ulaval.ca  
Center for Computational Genomics and Bioinformatics (CCGB),  
University of Minnesota, MN Id Identifier: MNS5286152 Clone ID:  
GQ0227\_B7\_K04 Clones available through: John Mackay, Ph. D.  
Professeur adjoint -Assistant professor EMAIL:  
jmackay@rsvs.ulaval.ca Centre de Recherche en Biologie Forestiere  
(Forest Biology Research Center) Universite Laval Quebec, Quebec  
CANADA G1K 7P4  
Plate: 7.B7 row: 04 column: K  
Seq primer: T7 primer.

FEATURES  
source

location/Qualifiers  
1. 928  
/organism="Picea glauca"  
/mol\_type="mRNA"  
/strain="Three trees Unknown"  
/db\_xref="taxon:3330"  
/clone="GQ0227\_B7\_K04"  
/sex="Hermaphrodite"  
/tissue\_type="Differentiating xylem from roots 1 cm in  
diameter or larger"  
/dev\_stage="Non-lignified xylem tissues from larger roots  
in early part of growing season (dune)."  
/lab\_host="E. coli DH10B cells"  
/clone\_lib="GQ022: ROOT XYLEM - mature trees"  
/note="Organ: Roots from 9 year old trees measuring  
approximately 4.5 m tall, and 10 cm in diameter. Vector:  
pBluescript II SK (+) XR; Site 1: Eco-R1; Site 2: Xho-1;  
cDNA was prepared from 5 mg of poly A+ selected RNA and  
was directionally ligated into the pBluescript II SK (+)  
XR vector (Stratagene), transformed by electroporation  
into DH10B cells (in vitro) for propagation"

ORIGIN

Query Match 24.3% Score 889.8; DB 7; Length 928;  
Best Local Similarity 96.0%; Pred. No. 3.2e-250;  
Matches 891; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

1915 TTTCATAGAGCTCGCCCCCTGACGAGCATCAAAAATCGACGTCGAAGCAGAGTG 1974  
1 TTTTATAGGCTCCGCCCTCGACGACATCAAAAATCGACGTCGAAGGCTCCCTGAGC 60

1975 GCGAAACCGACAGACTATTAAGATCCAGCGGTTTCCCTCGAAGGCTCCCTGAGC 2034  
61 GCGAAACCGACAGACTATTAAGATCCAGCGGTTTCCCTCGAAGGCTCCCTGAGC 120

2035 CTCTCGTTCGACCCCGCGCTTACCGGATACCTGTCCGCTTTCTCCCTGGGAAG 2094  
121 CTCTCGTTCGACCCCGCGCTTACCGGATACCTGTCCGCTTTCTCCCTGGGAAG 180

2095 CGTGGCGCTTTCATAGACTCAGCGTGTAGATCATGTTGCGTGTAGTGGTTCGTC 2154  
181 CGTGGCGCTTTCATAGACTCAGCGTGTAGATCATGTTGCGTGTAGTGGTTCGTC 240

2155 CAAGCTGGAGCTGTGACGAACCCCGTTCAAGCCGACGCGCTGCGCTTATCCGTTAA 2214  
241 CAAGCTGGAGCTGTGACGAACCCCGTTCAAGCCGACGCGCTGCGCTTATCCGTTAA 300

2215 CTATCGCTTGTAGTCAACCCCGTTAAGACACCATTTATTCGCACTGGACAGCACTGG 2274  
301 CTATCGCTTGTAGTCAACCCCGTTAAGACACCATTTATTCGCACTGGACAGCACTGG 360

2275 TAAAGAGATTAGAGAGCGAGGTAATGTAGCGGTCTTACAGAGTTCTTGAAGTGTGGCC 2334  
361 TAAAGAGATTAGAGAGCGAGGTAATGTAGCGGTCTTACAGAGTTCTTGAAGTGTGGCC 420

2335 TAACTAGGCTACACTAGAAGACAGTATTTGGTATCTGCGCTTGTGAAAGCACTTAC 2394

Db 421 TAACTACGGCTACACTAGAAAGATATTGTATCTGCGCTGTGTGAAGCCAGTTAC 480

Qy 2395 CTTGGGAAAAAGATTGTTACTCTGTATCCGGGAAACAAACACCGCTGGTAGCGGTG 2454

Db 481 CTTGGGAAAAAGATTGTTACTCTGTATCCGGGAAACAAACACCGCTGGTAGCGGTG 540

Qy 2455 TTTTGTGTGTGCAAGCAGCAGATTACGCGGAGAAAAAGGATCTCAAGAAATCCCTT 2514

Db 541 TTTTGTGTGTGCAAGCAGCAGATTACGCGGAGAAAAAGGATCTCAAGAAATCCCTT 600

Qy 2515 GATCTTTTCTACGGGGTCTGACGCTCAGTGGAAAGAAAACTCAGCTTAAGGATTTTGT 2574

Db 601 GATCTTTTCTACGGGGTCTGACGCTCAGTGGAAAGAAAACTCAGCTTAAGGATTTTGT 660

Qy 2575 CATAGATTTATCAAAAAAGATCTTCACTAGATCTTTTAAATTAAATGAAGTTTAA 2634

Db 661 CATAGATTTATCAAAAAAGATCTTCACTAGATCTTTTAAATTAAATGAAGTTTAA 720

Qy 2635 ATCAATCTAAAGTATATATGATTAACCTGGTCTGACAGTTACCAATGCTTAATCAGTGA 2694

Db 721 ATCAATCTAAAGTATATATGATTAACCTGGTCTGACAGTTACCAATGCTTAATCAGTGA 780

Qy 2695 GGCACTATCTCAGCGATCTGTATTTCTTCAATCCATAGTTGCTGACTCCCGTCT 2754

Db 781 GGCACTATCTCAGCGATCTGTATTTCTTCAATCCATAGTTGCTGACTCCCGTCT 840

Qy 2755 GTAATATCTACGATACGGAGGGCTTACCATCTGGCCCCAGTGTGCAATGATACCGG 2814

Db 841 GTNNNTATCTACGATACGGAGGGCTTACCATCTGNNNNNGTCTGCAATGATACNCG 900

Qy 2815 AGACCAAGCTCACCGGCTCCAGTTTA 2842

Db 901 AGACCAAGCTCACCGGCTCCAGTTTA 928

RESULT 8  
CB686151/c  
LOCUS  
DEFINITION  
Brn01b\_02008\_A 925 bp mRNA linear EST 09-APR-2003

Brn01b\_AAPC\_ECORC\_transgenic\_Brassica\_napus\_overexpressing\_BNCP17\_c  
onstitutively\_frost\_tolerant\_Brassica\_napus\_cDNA clone Brn01b\_02008,  
mRNA sequence.

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
CB686151  
CB686151.1 GI:29689876  
EST.

Brassica napus (rape)  
Brassica napus  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.

1 (bases 1 to 925)  
Singh,J., Allard,G., Tinker,N., Robert,L., Lacroix,C., De Moore,A.,  
Chagnon,J., Farah,S., Couroux,P. and Hattori,J.  
Expressed Sequence Tags from constitutively frost tolerant  
transgenic Brassica napus overexpressing BNCP17  
Unpublished (2002)  
Contact: Singh,J.A.

JOURNAL  
COMMENT  
Eastern Cereal and Oilseed Research Centre  
Agriculture and Agri-Food Canada  
KW Neelby Bldg., Central Experimental Farm, Ottawa, Ontario, K1A  
0G6, Canada

Tel: (613) 759-1662  
Fax: (613) 759-1701  
Email: singhj@agr.gc.ca.

FEATURES  
source  
Location/Qualifiers

1..925  
/organism="Brassica napus"  
/mol\_type="mRNA"  
/cultivar="Weiblar"  
/db\_xref="taxon:3708"  
/clone="Brn01b\_02008"  
/issue\_type="fourth leaf"  
/dev\_stage="3 weeks seedling grown at room temperature"

/clone\_id="Brn01b\_AAPC\_ECORC\_transgenic\_Brassica\_napus\_ove  
repressing\_BNCP17\_constitutively\_frost\_tolerant"  
/note="Vector: Bluescript SK+/XhoI-EcoRI; Site 1: EcoRI;  
Site 2: XhoI; Germinated in soil flats and seedlings grown  
for 3 weeks in a Conviron E-15 cabinet set at 20°C/16 hr  
light (250 Em-2sec-1) and 16°C/8 hr dark. Fourth leaves  
collected at 9 am and immediately frozen."

Query Match 24.1%; Score 881.4; DB 6; Length 925;  
Best Local Similarity 99.0%; Pred. No. 9.7e-248;  
Matches 890; Conservative 6; Mismatches 2; Indels 1; Gaps 1;

Qy 1493 GCGGCTTGGGTAATCATGTCTATAGCTGTTCTGTGTGAATTTGTTATCCGTCACA 1552

Db 898 GCGGCTTGGGTAATCATGTCTATAGCTGTTCTGTGTGAATTTGTTATCCGTCACA 840

Qy 1553 ATTCCACAAACATACAGAGCGGAGAGATTAAGGTAAAGCTGGGGTGCCCTAATGAGTG 1612

Db 839 ATTCCACAAACATACAGAGCGGAGAGATTAAGGTAAAGCTGGGGTGCCCTAATGAGTG 780

Qy 1613 AGCTAACTCAATTAATGGGCTTGCGCTCACTGCGCCCTTCCAGTGGGAAACCTGTG 1672

Db 779 AGCTAACTCAATTAATGGGCTTGCGCTCACTGCGCCCTTCCAGTGGGAAACCTGTG 720

Qy 1673 TGCCAGCTGCATTAATGATTCGCAACGCGCGGGAGAGCGGTTGCTATTGGCGGC 1732

Db 719 TGCCAGCTGCATTAATGATTCGCAACGCGCGGGAGAGCGGTTGCTATTGGCGGC 660

Qy 1733 TCTTCGCGCTTCTGCTGCATGACTGCTGCGCTGGCTCGTTGGCGCGGAGCGGTA 1792

Db 659 TCTTCGCGCTTCTGCTGCATGACTGCTGCGCTGGCTCGTTGGCGCGGAGCGGTA 600

Qy 1793 TCACTCACTCAAAAGCGGTAATACGGTTATCCACAGATCAGGGGATTAACGACAGAAAG 1852

Db 599 TCACTCACTCAAAAGCGGTAATACGGTTATCCACAGATCAGGGGATTAACGACAGAAAG 540

Qy 1853 AACATGTGACAAAGAGCCAGCAAAAGCCAGAGAACCTGTAAAGAGCCGCTTCTGCGG 1912

Db 539 AACATGTGACAAAGAGCCAGCAAAAGCCAGAGAACCTGTAAAGAGCCGCTTCTGCGG 480

Qy 1913 TTTTTCATATGGCTCGCGCCCGCCGAGAGATCAAAAATGAGCGCTCAAGTCAAGG 1972

Db 479 TTTTTCATATGGCTCGCGCCCGCCGAGAGATCAAAAATGAGCGCTCAAGTCAAGG 420

Qy 1973 TGGGAAACCCGACAGGACTATTAAGATACAGGCGTTTCCCTCGAAGCTCCCTGTG 2032

Db 419 TGGGAAACCCGACAGGACTATTAAGATACAGGCGTTTCCCTCGAAGCTCCCTGTG 360

Qy 2033 CGCTCTCTGTTCCGACCTGCGGCTTACCGGATCCTGTCCGCTTTCCTTGGGGA 2092

Db 359 CGCTCTCTGTTCCGACCTGCGGCTTACCGGATCCTGTCCGCTTTCCTTGGGGA 300

Qy 2093 AGGTGCGGCTTCTCACTACCTCAGCGGTGAGATCTCAGTTCCGTTGAGTGTGCG 2152

Db 299 AGGTGCGGCTTCTCACTACCTCAGCGGTGAGATCTCAGTTCCGTTGAGTGTGCG 240

Qy 2153 TCCAGCTGGGCTGTGACAGAACCCCGCTTACGCGGACCGCTGCGCTTATCCGCT 2212

Db 239 TCCAGCTGGGCTGTGACAGAACCCCGCTTACGCGGACCGCTGCGCTTATCCGCT 180

Qy 2213 AACTATGTTTGAATGTCAAACCGGTAAAGACAGACTTATCGCCATGAGACGCACT 2272

Db 179 AACTATGTTTGAATGTCAAACCGGTAAAGACAGACTTATCGCCATGAGACGCACT 120

Qy 2273 GGTAAACAGATTAGCAGAGGAGATGTTAGGGGCTTCAAGTCTTGAAGTGGTG 2332

Db 119 GGTAAACAGATTAGCAGAGGAGATGTTAGGGGCTTCAAGTCTTGAAGTGGTG 60

Qy 2333 CTTAACTACGGCTACATAAGAGAGATTTTGTATCTGCGCTGTGTAAGCCAGT 2391

Db 59 CTTAACTACGGCTACATAAGAGAGATTTTGTATCTGCGCTGTGTAAGCCAGT 1

RESULT 9  
BG838279/c 935 bp mRNA linear EST 25-MAY-2001

LOCUS BG838279.1 GI:14204601

DEFINITION Glycine claudestina cDNA clone Gc01\_10e07, mRNA sequence.

ACCESSION BG838279

VERSION BG838279.1

KEYWORDS EST

SOURCE Glycine claudestina

ORGANISM Glycine claudestina

REFERENCE 1 (bases 1 to 935)  
Singh, J.A., Farah, S., Chapados, J., Couroux, P., De Moors, A., Harris, L.J., Hattori, J.I., Ouellet, T., Robert, L.S., Sprott, D. and Tinker, N.A.  
Expressed Sequence Tags from Cold-Stressed Glycine claudestina Seedlings  
Unpublished (2001)  
Contract: Singh, J.A.  
Eastern Cereal and Oilseed Research Centre  
Agriculture and Agri-Food Canada  
KW Neaby Bldg., Central Experimental Farm, Ottawa, Ontario, K1A 0C6, Canada  
Tel: (613) 759-1662  
Fax: (613) 759-1701  
Email: singhja@agr.gc.ca

FEATURES  
Source  
1..935  
Location/Qualifiers  
/organism="Glycine claudestina"  
/mol\_type="mRNA"  
/cultivar="1035"  
/db\_xref="taxon:45687"  
/clone="Gc01\_10e07"  
/issue\_type="Leaves, stem"  
/clone\_id="Gc01\_AAPC\_ECORC\_cold\_stressed\_Glycine\_claudestina"  
/note="Vector: Bluescript SK+/XhoI-EcoRI; Site 1: EcoRI, Site 2: XhoI; Plants incubated at 2 degrees under 12 hours of light/day. Harvested after only 2-3 days of cold treatment. cDNA was prepared with the Uni-Zap cDNA kit from Stratagene. Eco RI adapters were linked followed by digest with Xho I/Eco RI and ligated to pbluescript."

ORIGIN

Query Match 23.8%; Score 870.2; DB 4; Length 935;  
Best Local Similarity 98.1%; Pred. No. 2e-244; Indels 1; Gaps 1;  
Matches 874; Conservative 14; Mismatches 2;

QY 2773 GGAGGGCTTACATCTGCGCCCAAGTGTGCAATGATACCGGAGACCCACGCTCACCGGC 2832

DB 928 GGGGGGGTTTACCATCTGCCCCCGAGKCTGCAATGATACCGGAGACCCACGCTCACCGGC 869

QY 2833 TCCAGATTATACGAAATAACCGAGCCAGCCGGAAGGCCGAGCCGAGAAAGTGTCTTC 2892

DB 868 YCCAGATTTATCTCGCATTAACCCAGCCAGCCGGAAGGCCGAGCCGAGAAAGTGTCTTC 809

QY 2893 AACTTATCTCGCTTCATCTCATCTATTAATTGTTGCGGGAAGCTAGAGTAAGTATTC 2952

DB 808 AACTTATCTCGCTTCATCTCATCTATTAATTGTTGCGGGAAGCTAGAGTAAGTATTC 749

QY 2953 GCCAGTTATAGTTGGCGGCAAGCTTGTTCCTCATCTGCTAAGGCAATCTGTGTCAAGCTC 3012

DB 748 GCCAGTTATAGTTGGCGGCAAGCTTGTTCCTCATCTAAGGCAATCTGTGTCAAGCTC 689

QY 3013 GTCTGTTGTTAGTGTTCATTCAGCTCCGGTTCACCAAGATTAAGCGCAATTAATGATC 3072

DB 688 GTCTGTTGTTAGTGTTCATTCAGCTCCGGTTCACCAAGATTAAGCGCAATTAATGATC 629

QY 3073 CCCCATGTTGTGCAAAAAAGCGTTAGCTCTTGGTGTCTCGATCTGTTGTCAAGATTA 3132

DB 628 CCCCATGTTGTGCAAAAAAGCGTTAGCTCTTGGTGTCTCGATCTGTTGTCAAGATTA 569

QY 3133 GTT-GGGCGGAGTTTATCACTCATGTGTTATGGCAGACATGCATATTTCTTACGTCA 3191

DB 568 GTTGGCGCGGAGTTTATCACTCATGTGTTATGGCAGACATGCATATTTCTTACGTCA 509

QY 3192 TGGCATCCGTAAGATCTTTTCTGTGATCTGTGATCTCAACCAAGTATCTGAGAT 3251

DB 508 TGGCATCCGTAAGATCTTTTCTGTGATCTGTGATCTCAACCAAGTATCTGAGAT 449

QY 3252 AGTGTATGCGCGCAGACGAGTGTCTTGGCCGGCGTCAATAGGGATTAATACCGGCCAC 3311

DB 448 AGTGTATGCGCGCAGACGAGTGTCTTGGCCGGCGTCAATAGGGATTAATACCGGCCAC 389

QY 3312 ATAGCAGAACTTTAAAGTCTCATCTATTTGAAAAAGTCTTGGGCGGCAAACTCTCA 3371

DB 388 ATAGCAGAACTTTAAAGTCTCATCTATTTGAAAAAGTCTTGGGCGGCAAACTCTCA 329

QY 3372 GGATCTTACCGCGTGTGAGATCCAGTGTGATGTAACCACTGTGCAACCCAACTGATCTT 3431

DB 328 GGATCTTACCGCGTGTGAGATCCAGTGTGATGTAACCACTGTGCAACCCAACTGATCTT 269

QY 3432 CAGCATCTTTTACTTTACACAGCGTTTCTGGGTGAGCAAAAAAGAGGCAAAATGCG 3491

DB 268 CAGCATCTTTTACTTTACACAGCGTTTCTGGGTGAGCAAAAAAGAGGCAAAATGCG 209

QY 3492 CAAAAAAGGAAATTAAGGCGGACACGGAATGTAATCTATCTCTCTTTTCAT 3551

DB 208 CAAAAAAGGAAATTAAGGCGGACACGGAATGTAATCTATCTCTCTTTTCAT 149

QY 3552 ATTATTTGAAGCAATTTATCAGGTTATTTGTCTCATGAGCGGATACATTTGAATGTTT 3611

DB 148 ATTATTTGAAGCAATTTATCAGGTTATTTGTCTCATGAGCGGATACATTTGAATGTTT 89

QY 3612 AGAAAAATTAACAAATAGGGGTTCCGGCACATTTCCCGAAAAAGTCCAC 3662

DB 88 AGAAAAATTAACAAATAGGGGTTCCGGCACATTTCCCGAAAAAGTCCAC 38

RESULT 10  
CF269652 1073 bp mRNA linear EST 13-AUG-2003  
LOCUS CF269652  
DEFINITION Fcyl1cold84 Fragilariopsis cylindrus SMART cDNA library (Clontech)  
Fragilariopsis cylindrus cDNA clone Antarctic 5', mRNA sequence.  
CF269652  
VERSION CF269652.1 GI:33631539  
KEYWORDS EST  
SOURCE Fragilariopsis cylindrus  
ORGANISM Fragilariopsis cylindrus  
REFERENCE 1 (bases 1 to 1073)  
Mock, T. and Valentin, K.  
EST analysis of freezing tolerance in the Antarctic diatom  
Fragilariopsis cylindrus: Detection of numerous cold adaption  
related genes and gene transfer events  
Unpublished (2003)  
CONTACT: Mock T  
Biological Oceanography  
Alfred-Wegener-Institute for Polar and Marine Research  
Am Handelshafen 12, D-27570 Bremerhaven, Germany  
Tel: +49 471 4831 1893  
Fax: +49 471 4831 1425  
Email: tmock@awi-bremerhaven.de  
sequence with unknown function  
PCR Primers  
FORWARD: 5'-lambdattt1p1EX2  
BACKWARD: 3'-lambdattt1p1EX2  
Seq primer: ctccggaagcgcccatcgtctgct.  
Location/Qualifiers  
1..1073  
/organism="Fragilariopsis cylindrus"



/mol\_type="mRNA"  
 /strain="Antarctic"  
 /db\_xref="taxon:186039"  
 /clone="Antarctic"  
 /clone\_1lb="Fragilariopsis cylindrus SMART cDNA library  
 (Clontech)"  
 /note="Vector: pTriplex2; total polyA was used for  
 first-strand synthesis with SMART IV oligos and CDS  
 III/3 PCR primer. Double strand cDNA synthesis was done by  
 LD PCR using the following program: 95°C for 5 min  
 denaturation and subsequent 20 cycles at 95°C (2min) and  
 68°C (6min). After SEI digestion the cDNA was  
 fractionated with CHROMA Spin-400 columns. These cDNAs  
 were ligated overnight into pTriplex2 vectors."

## ORIGIN

Query Match 23.6%; Score 865; DB 6; Length 1073;  
 Best Local Similarity 96.2%; Pred. No. 7e-243;  
 Matches 927; Conservative 0; Mismatches 30; Indels 7; Gaps 4;

1631 GCGTTGCGCTCACTGCGCGCTTTCCAGTCGGGAACTGTGTGCGACCTGCATTATGA 1690  
 1 GCTTCTTGCTCACTGCGCGCTTTCCAGTCGGGAACTGTGTGCGACCTGCATTATGA 60

1691 ATCGGCCAAGCGCGGGGAGAGGGGTTGGGATTTGGGCGCTTCCGCTTCCGCTC 1750  
 61 ATCGGCCAAGCGCGGGGAGAGGGGTTGGGATTTGGGCGCTTCCGCTTCCGCTC 120

1751 ACTGACTGCTGCGCTGCGCTGCTTCCGCGCGCGAGCGGTATCAGCTCACTCAAAAGCG 1810  
 121 ACTGACTGCTGCGCTGCGCTTCCGCGCGCGAGCGGTATCAGCTCACTCAAAAGCG 180

1811 GTAATACGCTTATCCAGAAATCAGGGGATTAACGAGAAAGAAATGTGACCAAAAGCG 1870  
 181 GTAATACGCTTATCCAGAAATCAGGGGATTAACGAGAAAGAAATGTGACCAAAAGCG 240

1871 CAGAAAAGCGCGAGAACCGTAAAGCGCGCTTGGCTGCTTTCATTAAGTCCGC 1930  
 241 CAGAAAAGCGCGAGAACCGTAAAGCGCGCTTGGCTGCTTTCATTAAGTCCGC 300

1931 CCCCCTGACGAGCATCAAAAATGAGCTCAAGTCAGAGGTGCGAAACCCGACAGGA 1990  
 301 CCCCCTGACGAGCATCAAAAATGAGCTCAAGTCAGAGGTGCGAAACCCGACAGGA 360

1991 CTATTAAGATACAGAGCGCTTCCCGCTGAAAGCTCCTGTCGCTCTCTGTTCCGAC 2050  
 361 CTATTAAGATACAGAGCGCTTCCCGCTGAAAGCTCCTGTCGCTCTCTGTTCCGAC 420

2051 CTGCGCTTACCGGATACCTGCTGCGCTTCTCCCTTGGGAAAGGTGCGCTTCTAT 2110  
 421 CTGCGCTTACCGGATACCTGCTGCGCTTCTCCCTTGGGAAAGGTGCGCTTCTAT 480

2111 AGCTCAGCTGATGATCTCAGTTGCGTGTAGTCTGCTCCAGACTGGGCTGTGTG 2170  
 481 AGCTCAGCTGATGATCTCAGTTGCGTGTAGTCTGCTCCAGACTGGGCTGTGTG 540

2171 CACGAAACCCCGCTTACGCGCGAGCGCTGCGCTTATCCGTAATCATGTTGAGTCC 2230  
 541 CACGAAACCCCGCTTACGCGCGAGCGCTGCGCTTATCCGTAATCATGTTGAGTCC 600

2231 AACCCGTAAGACAGACTTATCGGCACTGCGAGGACACTGTGTAACAGATTAGCAGA 2290  
 601 AACCCGTAAGACAGACTTATCGGCACTGCGAGGACACTGTGTAACAGATTAGCAGA 660

2291 GCGAGTATGATGAGCGGTGCTACAGAGTTCTTGAAGGTGGAGCT--AACTACGCTACA 2348  
 661 GCGAGTATGATGAGCGGTGCTACAGAGTTCTTGAAGGTGGAGCT--AACTACGCTACA 720

2349 CTAGAAGACA-GTATTTGTTGATCT-GCGCTCTGCTGAAGC--AGTTACCTTGGAAA 2403  
 721 CTAGAAGACAANGATTTGTTGATCTGCTGCTGAAGCAGANTTACNTTCGAAA 780

2404 AAGAGTTGATGCTTGTATCCGCAACCAACCGCTGTGAGCGGTGTTTTTGT 2463

DB 781 AAGAGTTGATGCTTGTATCCGCAACCAACCGCTGTGAGCGGTGTTTTTGT 840

2464 TTGCAAGCAGAGATTAACGCGCAGAAAAAGATCTCAAGAGATCTTTGATCTTTTC 2523  
 841 TTGCAAGCAGAGATTAACGCGCAGAAAAAGATCTCAAGAGATCTTTGATCTTTTC 900

2524 TACGGGCTTCAAGCTCAAGTGAACGAAACTCAGTTAAGGATTTTGTGATGACATT 2583  
 901 TACGGGCTTCAAGCTCAAGTGAACGAAACTCAGTTAAGGATTTTGTGATGACACA 960

2584 ATCA 2587  
 DB 961 ATAA 964

## RESULT 11

CL076016 885 bp DNA linear GSS 31-DEC-2003  
 LOCUS CH216-138F20.RM1.1 CH216 Xenopus tropicalis genomic clone  
 DEFINITION CH216-138F20, genomic survey sequence.  
 ACCESSION CL076016  
 VERSION CL076016.1 GI:40531929  
 KEYWORDS GSS.

## ORGANISM

Xenopus tropicalis (western clawed frog)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;  
 Xenopodidae; Xenopus; Silurana.

## REFERENCE

Kremetzki, C., Carter, J., McPherson, J., Warren, W., Graves, T.,

## AUTHORS

Mardis, E. and Wilson, R.

A physical map of the xenopus tropicalis genome

## JOURNAL

Unpublished (2003)

## COMMENT

Contact: Richard K Wilson

Genome Sequencing Center

Washington University School of Medicine

Email: submissions@wustl.edu

Insert Length: 175000 Std Error: 0.00

Seq primer: RM1 TACGACTCACTATGAGGAGA

Class: BAC ends

High quality sequence start: 11

High quality sequence stop: 810.

Location/Qualifiers

## FEATURES

## source

1..885

/organism="Xenopus tropicalis"

/mol\_type="genomic DNA"

/strain="Nigerian frog"

/db\_xref="taxon:8364"

/clone="CH216-138F20"

/sex="male"

/cell\_line="Stock 248 F7A2, inbred N7"

/clone\_1lb="CH216"

/note="Vector: pTARBAC2.1; CHORI-216 Xenopus tropicalis  
 BAC library"

## ORIGIN

Query Match 23.4%; Score 857.6; DB 9; Length 885;  
 Best Local Similarity 99.3%; Pred. No. 1e-240;  
 Matches 871; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

1535 AATTGTTATCCGCTCAACAATTCACAGCAATACAGAGCCGGAAGCATTAAGTAAAGCC 1594  
 5 AATTGTTATCCGCTCAACAATTCACAGCAATACAGAGCCGGAAGCATTAAGTAAAGCC 64

1595 TGGGTCCTTAATGAGGAGTAATCACTTAATTGCGTTGGCTACCTGCGCGCTTTC 1654  
 65 TGGGTCCTTAATGAGGAGTAATCACTTAATTGCGTTGGCTACCTGCGCGCTTTC 124

1655 CAGTCGGGAAACCTGTGCTCAGCTCAGCTCAATTAATGAATCGGCCAACCGCGGGAGAGGC 1714  
 125 CAGTCGGGAAACCTGTGCTCAGCTCAGCTCAATTAATGAATCGGCCAACCGCGGGAGAGGC 184

```

QY 1715 GATTGGGATTTGGGCGCTCTTCCGCTTCTCGCTCACTGACTGCGCTCGCTGCTT 1774
DB 185 GATTGGGATTTGGGCGCTCTTCCGCTTCTCGCTCACTGACTGCGCTCGCTGCTT 244
QY 1775 CGGCTGGCGGCGAGGCTATGACTCACTCAAGAGGCGGTATAGGTTATTCAGAGATCA 1834
DB 245 CGGCTGGCGGCGAGGCTATGACTCACTCAAGAGGCGGTATAGGTTATTCAGAGATCA 304
QY 1835 GGGGATTAACCGAGAGAAAGCATGTGAGCAAAAGGCCAGAGAAAGCCAGAAACGGTAA 1894
DB 305 GGGGATTAACCGAGAGAAAGCATGTGAGCAAAAGGCCAGAGAAAGCCAGAAACGGTAA 364
QY 1895 AAGGCGCGCTGGCTGGGCTTTTCCATAGGCTCCGCCCCCTGACGACATCAAAAAT 1954
DB 365 AAGGCGCGCTGGCTGGGCTTTTCCATAGGCTCCGCCCCCTGACGACATCAAAAAT 424
QY 1955 CGAGCTCAAGTCAAGAGTGGCGAAACCCGACAGACTTAAGATACAGGCGTTCC 2014
DB 425 CGAGCTCAAGTCAAGAGTGGCGAAACCCGACAGACTTAAGATACAGGCGTTCC 484
QY 2015 CTTGGAAGCTCCTCGTGGCTCTCCTGTTCCGACCCCTGCTTACCGATACCTGTC 2074
DB 485 CTTGGAAGCTCCTCGTGGCTCTCCTGTTCCGACCCCTGCTTACCGATACCTGTC 544
QY 2075 GCTTTCTCTCTTCCGGAAGCGTGGGCTTTCTCATAGTCAAGCTGATAGTCTAGT 2134
DB 545 GCTTTCTCTCTTCCGGAAGCGTGGGCTTTCTCATAGTCAAGCTGATAGTCTAGT 604
QY 2135 TCGGTGTAGTGTGCTTCCGCAAGCTGGGCTGTGTGCAAGAACCCCGTTCAGCCGAC 2194
DB 605 TCGGTGTAGTGTGCTTCCGCAAGCTGGGCTGTGTGCAAGAACCCCGTTCAGCCGAC 664
QY 2195 CGCTGCGCTTATCCGCTTATCTGCTTGTAGTCAAGCCGCTTACCGATACCTGTC 2254
DB 665 CGCTGCGCTTATCCGCTTATCTGCTTGTAGTCAAGCCGCTTACCGATACCTGTC 724
QY 2255 CCACTGGAGAGCAAGCTGTTAAGAGATTGACAGAGAGATGTTAGGCGGTCTACA 2314
DB 725 CCACTGGAGAGCAAGCTGTTAAGAGATTGACAGAGAGATGTTAGGCGGTCTACA 784
QY 2315 GAGTTCCTTGAAGTGTGCTTACTACTAGCGCTACACTGAAAGAGATTTGGTATTCG 2374
DB 785 GAGTTCCTTGAAGTGTGCTTACTACTAGCGCTACACTGAAAGAGATTTGGTATTCG 844
QY 2375 GCTGCTGAGAGCAAGTACTTCTGAAAAGAGTTG 2411
DB 845 GCTGCTGAGAG-CAAGTACTTCTGAAAAGAGTTG 880

```

```

RESULT 12
AL042026/c 841 bp mRNA linear EST 04-SHP-2003
LOCUS DKFZp434E111.k1.434 (synonym: hhes3) Homo sapiens cDNA clone
DEFINITION DKFZp434E111.5, mRNA sequence.
ACCESSION AL042026
VERSION AL042026.1 GI:5421372
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 841)
Pousetka,A., Klein,M., Mewes,H.W., Gassenhuber,J. and Wiemann,S.
EST (Pousetka, et al.)
Unpublished (1999)
CONTACT: MIPS
MIPS
Ingolstaedter Landstr.1, D-85764 Neuberg, Germany
This is the 5' sequence of the clone insert
Cloned from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ), Email s.wiemann@dkfz-heidelberg.de;
Heidelberg/Germany) within the cDNA sequencing consortium of the

```

German Genome Project.  
No SI sequence available.  
This clone (DKFZp434E111) is available at the RZPD in Berlin.  
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059  
Berlin-Charlottenburg, GERMANY; Email: clone@rzd.de.

FEATURES  
Location/Qualifiers

```

1..841
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp434E111"
/issue_type="testis"
/dev_stage="adult"
/lab_host="DH10B"
/clone_id="434 (synonym: hhes3)"
/notes="Vector: pSPORT1; Site_1: NotI; Site_2: SalI"

```

## ORIGIN

Query Match 23.0%; Score 841; DB 1; Length 841;  
Best Local Similarity 100.0%; Pred. No. 7.8e-236;  
Matches 841; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 2788 TGGCCCAAGTGTGCAATGATACCGGAGACCAAGCTCAAGGCTCCAGATTATACG 2847
DB 841 TGGCCCAAGTGTGCAATGATACCGGAGACCAAGCTCAAGGCTCCAGATTATACG 782
QY 2848 AATTAACCAAGCCAGCGGAGAGGCGGAGCGCAAGAGTGTCTGCAACTTTATCCGCTC 2907
DB 781 AATTAACCAAGCCAGCGGAGAGGCGGAGCGCAAGAGTGTCTGCAACTTTATCCGCTC 722
QY 2908 CATCCAGTCTATTATTATGTTCCGGAAGCTAGAGTAAGTATGTCGACGTTATAGTT 2967
DB 721 CATCCAGTCTATTATTATGTTCCGGAAGCTAGAGTAAGTATGTCGACGTTATAGTT 662
QY 2968 GCGCAAGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTG 3027
DB 661 GCGCAAGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTG 602
QY 3028 TTCAATTCAGCTCCGAGTCCCAAGCATCAAGCCAGTTCATGATCCGCCATGTTGCA 3087
DB 601 TTCAATTCAGCTCCGAGTCCCAAGCATCAAGCCAGTTCATGATCCGCCATGTTGCA 542
QY 3088 AAAAGCGGTATAGCTCTTCCGCTCCGATCGTTGTGCAAGATGTTGCGCGAGTGT 3147
DB 541 AAAAGCGGTATAGCTCTTCCGCTCCGATCGTTGTGCAAGATGTTGCGCGAGTGT 482
QY 3148 ATCACTCATGTTATGAGCAGACATGATATTTCTTCACTGTCAAGCCATCCGTAAGATG 3207
DB 481 ATCACTCATGTTATGAGCAGACATGATATTTCTTCACTGTCAAGCCATCCGTAAGATG 422
QY 3208 CTTTCTGTGACTGTGAGTACTCAACCAAGTCTTCTGAGATAGTGTATGCGCGAGC 3267
DB 421 CTTTCTGTGACTGTGAGTACTCAACCAAGTCTTCTGAGATAGTGTATGCGCGAGC 362
QY 3268 GAGTGTCTTGTGCGCGCGCAATACGGGATTAATCCGGGCACTAGCAAGAACTTTAA 3327
DB 361 GAGTGTCTTGTGCGCGCGCAATACGGGATTAATCCGGGCACTAGCAAGAACTTTAA 302
QY 3328 AGTGTCTATTTGAAAAAGTCTTCTGCGGCGAAAACTCTCAAGATCTTACCGCTGT 3387
DB 301 AGTGTCTATTTGAAAAAGTCTTCTGCGGCGAAAACTCTCAAGATCTTACCGCTGT 242
QY 3388 GAGATCCAGTGTGAGTAAACCACTGCGGCAACCCAGCTGATCTTCAAGATCTTTACTTT 3447
DB 241 GAGATCCAGTGTGAGTAAACCACTGCGGCAACCCAGCTGATCTTCAAGATCTTTACTTT 182
QY 3448 CACCAAGCTTTTGTGAGTGAAGAAAAAGAGGCAAAATCCGCAAAAAGGAATTAAG 3507
DB 181 CACCAAGCTTTTGTGAGTGAAGAAAAAGAGGCAAAATCCGCAAAAAGGAATTAAG 122
QY 3508 GCGGACAGGAAATGTTAAATGCTCATGCTTCTTTTCAATATTAATGAAGCATTTA 3567
DB 121 GCGGACAGGAAATGTTAAATGCTCATGCTTCTTTTCAATATTAATGAAGCATTTA 62

```

QY 3568 TCAGGTTATGCTCATGACGGATACATATTTGATGATTTAGAAAATAAACAAT 3627  
DB 61 TCAGGTTATGCTCATGACGGATACATATTTGATGATTTAGAAAATAAACAAT 2  
QY 3628 A 3628  
DB 1 A 1  
RESULT 13  
BM438950/c 854 bp mRNA linear EST 31-JAN-2002  
LOCUS Ipvtr00491 Liver cDNA library Ictalurus punctatus cDNA 5', mRNA  
DEFINITION  
ACCESSION BM438950  
VERSION BM438950.1 GI:18460672  
KEYWORDS EST.  
SOURCE Ictalurus punctatus (channel catfish)  
ORGANISM Ictalurus punctatus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Actinopterygii; Neopterygii; Teleostei; Osteichthyes; Siluriformes;  
Ictaluridae; Ictalurus.  
1 (bases 1 to 854)  
Feng J., Kucukas H., Kocabas A., Li P. and Liu Z.  
Transcriptome of channel catfish (Ictalurus punctatus): initial  
analysis of expressed sequence tags from the liver  
unpublished (2002)  
JOURNAL Contact: Liu Z.  
The Fish Molecular Genetics and Biotechnology Laboratory,  
Department of Fisheries and Allied Aquacultures and Program of Cell  
and Molecular Biosciences  
Auburn University  
203 Swingle Hall, Auburn University, Auburn, AL 36849, USA  
Tel: 334 844 4054  
Fax: 334 844 9208  
Email: zliu@aesag.auburn.edu  
Seq primer: M13 Reverse.  
Location/Qualifiers  
Source 1. 854  
/organism="Ictalurus punctatus"  
/mol\_type="mRNA"  
/db\_xref="taxon:7998"  
/clone\_lib="Liver cDNA library"  
/note="Organ: Liver; Vector: pSport1; Site\_1: Not1;  
Site\_2: SalI"

ORIGIN  
Query Match 22.5%; Score 824.6; DB 4; Length 854;  
Beet Local Similarity 99.3%; Pred. No. 5.4e-231;  
Matches 849; Conservative 0; Mismatches 4; Indels 2; Gaps 2;  
QY 2636 TCAATCTAAGTAAATAGTAACTGTCTGACAGTACCAATGCTTATCAGTAC 2695  
DB 854 TCAATCTAAGTAAATAGTAACTGTCTGACAGTACCAATGCTTATCAGTAC 795  
QY 2696 GCACCTATCTCAGGAGTCTGTAATTTCTTCATTCATGATGTTGCCCTGCTG 2755  
DB 794 GCACCTATCTCAGGAGTCTGTAATTTCTTCATTCATGATGTTGCCCTGCTG 735  
QY 2756 TAGATTAAGTAAAGGAGGCTTACATCTGCGCCCAATGCTGCAATATACCGGA 2815  
DB 734 TAGATTAAGTAAAGGAGGCTTACATCTGCGCCCAATGCTGCAATATACCGGA 675  
QY 2816 GACCCAGCTCAGCGGCTCAGATTTACAGCAATTAACAGCCAGCGGAAGGCGAG 2875  
DB 674 GACCCAGCTCAGCGGCTCAGATTTACAGCAATTAACAGCCAGCGGAAGGCGAG 615  
QY 2876 CGCGAAGTGTCTCGCACTTTATCGCGCTTCATCCAGTCTATTATTTGCGCGGA 2935  
DB 614 CGCGAAGTGTCTCGCACTTTATCCGCTTCATCCAGTCTATTATTTGCGCGGA 555  
QY 2936 GCTAGAGTAAAGTGTGCGCAGTTAATAGTTGGCAACGTTTGCATTGCTACAGGC 2995

DB 554 GCTAGAGTAAAGTGTGCGCAGTTAATAGTTGGCAACGTTTGCATTGCTACAGGC 495  
QY 2996 ATCGTGTGTACGCTCGTCTGTTGATAGCTTCATTCAGCTCCGTTCCCAAGATCA 3055  
DB 494 ATCGTGTGTACGCTCGTCTGTTGATAGCTTCATTCAGCTCCGTTCCCAAGATCA 435  
QY 3056 AGCGAGTTCATGATATCCCATGTTGTGCAAAAAGCGTTAGCTCTTGGTCTCCG 3115  
DB 434 AGCGAGTTCATGATATCCCATGTTGTGCAAAAAGCGTTAGCTCTTGGTCTCCG 376  
QY 3116 ATCGTGTGCAAGTAAAGTGTGCGCAGTGTATACATGATGTTAAGGACAGCTGAT 3175  
DB 375 ATCGTGTGCAAGTAAAGTGTGCGCAGTGTATACATGATGTTAAGGACAGCTGAT 316  
QY 3176 AATTCCTTACTGTCATGATCCGATCCGTAAGATGCTTTCTGTAGCTGTGATCTCAAC 3235  
DB 315 AATTCCTTACTGTCATGATCCGATCCGTAAGATGCTTTCTGTAGCTGTGATCTCAAC 256  
QY 3236 AAGTCATTTGGAAGTATGTTATGCGGACCGAGTGTCTTGC CGGCTCAATACG 3295  
DB 255 AAGTCATTTGGAAGTATGTTATGCGGACCGAGTGTCTTGC CGGCTCAATACG 196  
QY 3296 GATAATACCGGCGCACATGACGAACTTTAAAGTGTCTCATTTGAAAAAGTTCTTG 3355  
DB 195 GATAATACCGGCGCACATGACGAACTTTAAAGTGTCTCATTTGAAAAAGTTCTTG 136  
QY 3356 GGGGAAACCTCTCAAGATCTTACCGCTGTGATGATCCAGTGTGATGA-CCACTCG 3414  
DB 135 GGGGAAACCTCTCAAGATCTTACCGCTGTGATGATCCAGTGTGATGAAGCCACTCG 76  
QY 3415 TGCAACCACTGATCTTCAAGATCTTTTACTTTTCAACGAGGTTTGGGTGAGCAAAAC 3474  
DB 75 TGCAACCACTGATCTTCAAGATCTTTTACTTTTCAACGAGGTTTGGGTGAGCAAAAC 16  
QY 3475 AGGAAGGCAAAATGC 3489  
DB 15 AGGAAGGCAAAATGC 1  
RESULT 14  
AG332951 1169 bp DNA linear GSS 02-JUN-2004  
LOCUS Mus musculus molossinus DNA, clone:MSHg01-123f02.T7, genomic survey  
DEFINITION  
ACCESSION AG332951  
VERSION AG332951.1 GI:47906261  
KEYWORDS GSS.  
SOURCE Mus musculus molossinus  
ORGANISM Mus musculus molossinus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE  
1 Hattori M., Toyoda A., Noguchi H., Kojima T. and Sakaki Y.  
BAC end Sequences of Library MSHg01  
JOURNAL Unpublished  
AUTHORS Hattori M., Toyoda A., Noguchi H., Kojima T. and Sakaki Y.  
TITLE Direct Submision  
JOURNAL Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical  
and Chemical Research (RIKEN), Genomic Sciences Center (GSC),  
1-7-22 Suehiro-Chou, Tsukuba, Ibaraki, Japan, Yokohama, Kanagawa 230-0045, Japan  
(E-mail:hattori@gsc.riken.jp, URL:http://hgp.gsc.riken.go.jp/,  
Tel:81-45-503-9111, Fax:81-45-503-9170)  
Clones are derived from the mouse BAC library MSHg01.  
Library availability, please contact Kunya Abe (abe@erc.riken.jp).  
Tsukuba Institute, Bio Resource Center.  
The Institute of Physical and Chemical Research (RIKEN) 3-1-1  
Koyadai, Tsukuba, 305-0074 Japan  
phone: 81-298-36-9189, fax: 81-298-36-9199  
e-mail: abe@erc.riken.jp  
PRIMERS  
Sequencing : T7

LIBRARY : DBACE3.6  
 Vector : Ecoli  
 R.Site 1 : Ecoli  
 R.Site 2 : Ecoli  
 Location/Qualifiers  
 1..1169  
 /organism="Mus musculus molossinus"  
 /mol\_type="genomic DNA"  
 /sub\_species="molossinus"  
 /db\_xref="taxon:57486"  
 /clone="MSM901-123F02.17"  
 /sex="male"  
 /issue\_type="mixture of kidney and spleen"  
 /clone\_id="MSM901 Mouse Male BAC library"

## ORIGIN

Query Match 22.5%; Score 824; DB 9; Length 1169;  
 Best Local Similarity 89.3%; Pred. No. 8.8e-231;  
 Matches 922; Conservative 0; Mismatches 106; Indels 5; Gaps 3;

```

QY 1486 CCGCGGGGCGGCTTGGCGTAATCATGTCATAGCTTCTCGTGTGAATTTGATCC 1545
DB 130 CAGGCATGCAAGCTTGGCGTAATCCGTGCTTACGTGTTCTGTGTGAATTTGATCC 189
QY 1546 GCTCACAATTCCACACAATACAGAGCCGAAACATAAAGTGAAGCTGGGCTGCTTA 1605
DB 190 CTTCCCAATTCCTCCCAACATACAGAGCCGAAACATAAAGTGAAGCTGGGCTGCTTA 249
QY 1606 ATGAGTGAAGCTACTACATTAATTCGCTTGGCGTCACTGCCCGTTCCAGTGGGAAA 1665
DB 250 ATGAGTGAAGCTACTACATTAATTCGCTTGGCGTCACTGCCCGTTCCAGTGGGAAA 309
QY 1666 CCGTGTGCGCAAGCTGCAATTAATGATGCGCAACGCGGAGAGGCGGTTTGGCGAT 1725
DB 310 CCGTGTGCGCAAGCTGCAATTAATGATGCGCAACGCGGAGAGGCGGTTTGGCGAT 369
QY 1726 TGGGCGCTCTTCCGCTTCCGCTCACTGACGCTGCGTCCGCTGCGCTGCGGCG 1785
DB 370 TGGGCGCTCTTCCGCTTCCGCTCACTGACGCTGCGTCCGCTGCGCTGCGGCG 429
QY 1786 AGCGGTATCAAGCTCACTCAAAAGCGGTAAATCGGTTATCCAGAAATCAGGGGATAAC 1845
DB 430 AGCGGTATCAAGCTCACTCAAAAGCGGTAAATCGGTTATCCAGAAATCAGGGGATAAC 489
QY 1846 AGGAAGAACAATGTGAGCAAAAGCCGCAAAAGCCGCAAGACCGTAAAGCCGCTT 1905
DB 490 AGGAAGAACAATGTGAGCAAAAGCCGCAAAAGCCGCAAGACCGTAAAGCCGCTT 549
QY 1906 GCTGCGCTTTTTCATAGGCTCCGCCCTCCCTGACGACATCAAAAATCGAGCTCAG 1965
DB 550 GCTGCGCTTTTTCATAGGCTCCGCCCTCCCTGACGACATCAAAAATCGAGCTCAG 609
QY 1966 TCAGAGGTGAGGAACCCGACAGAGATTAATAAAT7ACAGAGCTTTTCCCTGGAAGCTC 2025
DB 610 TCAGAGGTGAGGAACCCGACAGAGATTAATAAAT7ACAGAGCTTTTCCCTGGAAGCTC 669
QY 2026 CTTGTCGCTCTTCTGTTCCGACCTGCGCTTACCGGATACCTGTCCGCTTTCTCCC 2085
DB 670 CTTGTCGCTCTTCTGTTCCGACCTGCGCTTACCGGATACCTGTCCGCTTTCTCCC 729
QY 2086 TTGCGGAAGCGTGGCGCTTCTCATAGCTCAAGCTGTAGTATCTCAAGTTCGAGTAGAGT 2145
DB 730 CTTGCGGAAGCGTGGCGCTTCTCATAGCTCAAGCTGTAGTATCTCAAGTTCGAGTAGAGT 789
QY 2146 CGTTGCTCTCAAGCTGGCTGTGTGCAAGAACCCCGCTTACGCGGACCGCTGCGCTT 2205
DB 790 CGTTGCTCTCAAGCTGGCTGTGTGCAAGAACCCCGCTTACGCGGACCGCTGCGCTT 849
QY 2206 ATCCGTAATCTAGCTTGAAGTCAACCCGGAAGACACATTAATGCGCATGGAGC 2265
DB 850 ATCCGTAATCTAGCTTGAAGTCAACCCGGAAGACACATTAATGCGCATGGAGC 909
QY 2266 AGCCATGTGTAAAGAGATTAGCAGAGCGAGGTATGTAGGCGGTCTACAGAGTTCTTGA 2325

```

```

DB 910 ACCCACTGTAAACGAGATTAGCAGAACCAAGTATGTAGCCCGCTAC-CAGTTCTGTAA 968
QY 2326 GTGTGCGCTTAACAGAGCTTACACTGAGAGACAGATTTGTATTTGGCTTGTGTAA 2385
DB 969 GGGTGGGCCCAACTACCGCTACCTTAAABAAAACATATTTGGATTCGCCCTCCGAAA 1028
QY 2386 GCCAGTTACCTTGGAAAAAGATTGTAGTACTTGTATCCGGCAACAAACGACCGCTG 2445
DB 1029 CCCCTTACCTTCCG--AAAAAATTGTACTTGTACCTTGAACCCCAA--AAACCCCTGT 1084
QY 2446 TAGCGGTGTTTTTTTGTGTCAGAGACAGATTAAGCCGCAAAAAAAGATCTCAAG 2505
DB 1085 AAACGGGGGTTTTTTTTTTTGTGCCCCACCCCAAAATCCCCCAAAAAAATCCCAAAA 1144
QY 2506 AGATCTTGTATC 2518
DB 1145 TCTTTTATTTTTC 1157

```

RESULT 15  
 CL021193  
 LOCUS  
 DEFINITION  
 CH216-8A16.RM.1 CH216 Xenopus tropicalis genomic clone CH216-8A16,  
 genomic survey sequence.  
 CL021193  
 ACCESSION  
 CL021193  
 VERSION  
 CL021193.1 GI:40463006  
 KEYWORDS  
 GSS.  
 SOURCE  
 Xenopus tropicalis (western clawed frog)  
 ORGANISM  
 Xenopus tropicalis  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;  
 Xenopodinae; Xenopus; Silurana.  
 1 (bases 1 to 1025)  
 AUTHORS  
 Kremitzki, C., Carter, J., McPherson, J., Warren, W., Graves, T.,  
 Mardis, E. and Wilson, R.  
 TITLE  
 A physical map of the xenopus tropicalis genome  
 JOURNAL  
 Unpublished (2003)  
 COMMENT  
 Contact: Richard K Wilson  
 Genome Sequencing Center  
 Washington University School of Medicine  
 Email: submissions@wustl.edu  
 Insert length: 175000 Std Error: 0.00  
 Seq primer: RM1 TACGACTCACTATAGGAGGA  
 Clones: BAC ends  
 High quality sequence start: 72  
 High quality sequence stop: 837.  
 Location/Qualifiers  
 1..1025  
 /organism="Xenopus tropicalis"  
 /mol\_type="genomic DNA"  
 /strain="Nigerian frog"  
 /db\_xref="taxon:8364"  
 /clone="CH216-8A16"  
 /sex="male"  
 /cell\_line="Stock 248 F7A2, inbred N7"  
 /clone\_id="CH216"  
 /note="Vector: pTARBAC2.1; CHORI-216 Xenopus tropicalis  
 BAC library"

## FEATURES

source  
 1..1025  
 /organism="Xenopus tropicalis"  
 /mol\_type="genomic DNA"  
 /strain="Nigerian frog"  
 /db\_xref="taxon:8364"  
 /clone="CH216-8A16"  
 /sex="male"  
 /cell\_line="Stock 248 F7A2, inbred N7"  
 /clone\_id="CH216"  
 /note="Vector: pTARBAC2.1; CHORI-216 Xenopus tropicalis  
 BAC library"

## ORIGIN

Query Match 22.4%; Score 821.4; DB 9; Length 1025;  
 Best Local Similarity 95.4%; Pred. No. 5e-230;  
 Matches 900; Conservative 0; Mismatches 36; Indels 7; Gaps 5;

```

QY 1489 GCGGCGCGCTTGGCGTAATCATGTCATAG-CTGTTCTGTGTGAATTTGATC-CG 1546
DB 84 GATGCAATCTTGGGTATCATGTCTAGTCTGTTCCGTGTGAATTTGATTCG 143
QY 1547 CTCACAATTCCACAACATACAGAGCGGAAGCATTAAGTGAAGCTGGGCTGCTAA 1606
DB 144 CTCACAATTCCACAACATCTTCTACCGGAAGCATTAAGTGAAGCGGGGCTGCTAA 203

```



**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 25, 2005, 23:20:55 ; Search time 1735.03 Seconds  
(without alignments)  
11079.575 Million cell updates/sec

Title: US-10-826-523-41

Perfect score: 3662

Sequence: 1 ctaaatgttaagcgcttaata.....attccccgaagtcgcac 3662

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 413486 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_23Sep04:\*

1: geneseqn1980s:\*

2: geneseqn1990s:\*

3: geneseqn2000s:\*

4: geneseqn2001as:\*

5: geneseqn2001bs:\*

6: geneseqn2002as:\*

7: geneseqn2002bs:\*

8: geneseqn2003as:\*

9: geneseqn2003bs:\*

10: geneseqn2003cs:\*

11: geneseqn2003ds:\*

12: geneseqn2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3648.4	99.6	3661	10	ACA55354 Transform
2	2665.4	72.8	3637	11	ADM68463 Variant g
3	2665.4	72.8	3637	12	ADP26616 Green flu
4	2664.8	72.8	3637	8	ABX14571 Plasmid p
5	2664.8	72.8	3637	11	ADM68448 Cycle 3 g
6	2664.8	72.5	3637	12	ADP26601 Green flu
7	2664.4	72.5	3637	11	ADM684570 Plasmid p
8	2664.4	72.5	3637	11	ADM68462 Wild type
9	2664.4	72.5	3637	12	ADP26615 Green flu
10	2647.8	72.3	3637	10	AD126329 Novel end
11	2635.8	72.0	3681	2	AAQ13578 Plasmid p
12	2622	71.6	5314	2	AAQ13576 Plasmid p
13	2620.6	71.6	3699	2	AAV14340 Plasmid p
14	2609.4	71.3	4001	6	AAD27066 Plasmid p
15	2484.6	67.8	3774	6	AAD27062 Plasmid p
16	2474.2	67.6	5277	3	AA88110 Plasmid p
17	2464.6	67.3	3908	6	ABT08165 Recombina
18	2451.8	67.0	3928	8	ABT14478 HCV envel
19	2451.8	67.0	3928	8	AAD50628 PBKS-E26H
20	2448.8	66.9	3956	2	ABO84236 Vector pb
21	2448.8	66.9	3956	2	AAV64258 Plasmid p

## ALIGNMENTS

22	2446.8	66.8	3715	6	AAD27064	Aad27064 Plasmid T
23	2446.8	66.8	3927	6	ABT08166	ABT08166 Recombina
24	2314.8	63.2	4088	2	AAV64255	AAV64255 Plasmid p
25	2305.2	62.9	4102	2	AAV64257	AAV64257 Plasmid p
26	2279.2	62.2	4119	6	AAD40772	Aad40772 Plasmid p
27	2271.2	62.0	4059	12	ADN97131	Adn97131 Mammalian
28	2230.6	60.9	5250	10	AAD49957	Aad49957 TMR (Tran
29	2224.6	60.7	2958	2	AAZ22250	Aaz22250 Nucleotid
30	2216.2	60.5	4205	4	AAD09088	Aad09088 PEAG658 p
31	2216.2	60.5	4205	6	AAK99713	Aak99713 DNA of pl
32	2215.4	60.5	4172	2	AAK19901	Aak19901 Plasmid p
33	2213	60.4	4205	4	AAD09087	Aad09087 PEAG657 p
34	2213	60.4	4205	6	AAK99700	Aak99700 DNA of pl
35	2202.6	60.1	4454	4	AA806386	AA806386 Vector pg
36	2172.8	59.3	7102	2	AAK13836	Aak13836 Nucleotid
37	2172.8	59.3	7102	6	ABK13259	Abk13259 Yeast pla
38	2172.8	59.3	7333	2	AAK1837	Abk1837 Nucleotid
39	2172.8	59.3	7333	6	ABK13260	Abk13260 Yeast pla
40	2172.8	59.3	7803	6	ABT08177	ABT08177 Recombina
41	2172.8	59.3	8167	6	ABT08178	ABT08178 Recombina
42	2172.6	59.3	10600	11	ADM68432	Adm68432 Celery CE
43	2172.6	59.3	10600	12	ADP26585	Adp26585 TMV CEL I
44	2172.6	59.3	10624	11	ADM68433	Adm68433 Celery CE
45	2172.6	59.3	10624	12	ADP26586	Adp26586 TMV CEL I

## RESULT 1

ACA55354	ACA55354 standard; DNA; 3661 BP.
ID	ACA55354 standard; DNA; 3661 BP.
AC	ACA55354;
DT	06-JUN-2003 (first entry)
DE	Transformation vector piggyBAC related plasmid pYL-Bac.
KW	PiggyBac; transposon; eukaryotic transformation vector; ds;
KW	transformed cell; transformed embryo; transgenic; plasmid; cyclic;
KW	circular.
OS	Synthetic.
XX	
XX	US2002173634-A1.
XX	
PD	21-NOV-2002.
XX	
XX	30-OCT-2001; 2001US-00001189.
XX	
XX	31-OCT-2000; 2000US-0244677P.
XX	
PR	01-NOV-2000; 2000US-0244984P.
XX	
PA	(FRAS/) FRASER M J.
PA	(LIXX/) LI X.
PA	(BEAM/) BEAM T.
PA	(HUA/) HUA-VAN A.
XX	
PI	Fraser MJ, Li X, . Beam T, Hua-Van A;
XX	
DR	WPI; 2003-352597/33.
XX	
PT	New DNA molecule in the transposon piggyBac, useful for transferring
PT	genes into host cells or embryos for transforming the cells of embryos
PT	that can be used in making transgenic organisms.
XX	
PS	Example 3, Fig 3(C2); 151pp; English.
XX	
CC	The invention describes a DNA molecule comprising at least 163
CC	consecutive nucleotide base pairs of the 3' terminal region beginning at
CC	the 3' terminal base pair, and at least 125 consecutive nucleotide base
CC	pairs of the 5' terminal region beginning at the 5' terminal base pair, of

CC the piggyBac molecule. The region extends from the restriction site SacI  
CC to the end of the piggyBac molecule. The DNA molecule in the transposon  
CC is useful for transfecting genes into host cells or embryos for  
CC piggyBac is useful for transfecting genes into host cells or embryos are  
CC transforming the cells of embryos. The transformed cells or embryos are  
CC useful for developing or making transgenic organisms. This sequence  
CC represents plasmid used in the creation of minimal sequence eukaryotic  
CC transformation vector piggyBac

XX Sequence 3661 BP; 910 A; 910 C; 957 G; 884 T; 0 U; 0 Other;

Query Match 99.6%; Score 3648.4; DB 10; Length 3661;

Best Local Similarity 99.9%; Pred. No. 0; Mismatches 1; Indels 1; Gaps 1;

Matches 3660; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 CTAATTTGTAAGCGTTAATTTGTTAAATTCGCTTAATTTGTTAAATTCGCTC 60  
DB 1 CTAATTTGTAAGCGTTAATTTGTTAAATTCGCTTAATTTGTTAAATTCGCTC 60  
QY 61 ATTTTAAACCAATAGCGCAATTCGCAAAATCCCTTAATTAATCAAAAGATAGACCA 120  
DB 61 ATTTTAAACCAATAGCGCG-ATTCGCAAAATCCCTTAATTAATCAAAAGATAGACCA 119  
QY 121 GATAGGTTGAGTGTGTTCCAGTTTGAACAAGTCACTATTAAAGACGTGACTC 180  
DB 121 GATAGGTTGAGTGTGTTCCAGTTTGAACAAGTCACTATTAAAGACGTGACTC 179  
QY 181 CAACGTCAAGGCGCAAAACCGTCTATCAGGCGGATGAGCCACTAGTGAACATCAC 240  
DB 180 CAACGTCAAGGCGCAAAACCGTCTATCAGGCGGATGAGCCACTAGTGAACATCAC 239  
QY 241 CTATCAATTTTGGGGTTCAGGTCGCTTAAGCACTAATCGGACCTTAAGGAG 300  
DB 241 CTATCAATTTTGGGGTTCAGGTCGCTTAAGCACTAATCGGACCTTAAGGAG 299  
QY 240 CTATCAATTTTGGGGTTCAGGTCGCTTAAGCACTAATCGGACCTTAAGGAG 299  
DB 301 CCCCCATTAGAGCTTGAAGCGGAAAGCGGCGCACTGCGAGAAAGAGAGAGAA 360  
QY 300 CCCCCATTAGAGCTTGAAGCGGAAAGCGGCGCACTGCGAGAAAGAGAGAGAA 359  
DB 361 AGGAAAGAGCGGCGCTAGGCGGCTGCAAGTGAAGCTGCAAGCTGCGTAAACAC 420  
QY 360 AGGAAAGAGCGGCGCTAGGCGGCTGCAAGTGAAGCTGCAAGCTGCGTAAACAC 419  
DB 421 CACACCGCGCGCTTAATGCGCGCTAAGCGGCGCTTCCATTCGCTTAAGGCTG 480  
QY 420 CACACCGCGCGCTTAATGCGCGCTAAGCGGCGCTTCCATTCGCTTAAGGCTG 479  
DB 481 CAACTGTTGGGAAGGCGGATCGGTGCGGCGCTTCCATTCGCTTAAGGCTG 540  
QY 480 CAACTGTTGGGAAGGCGGATCGGTGCGGCGCTTCCATTCGCTTAAGGCTG 539  
DB 541 GGAATGTCCTCAAGGCGATTAAGTGGTGAAGCGCGGCTTCCATTCGCTTAAG 600  
QY 540 GGAATGTCCTCAAGGCGATTAAGTGGTGAAGCGCGGCTTCCATTCGCTTAAG 599  
DB 601 TAAACGACGCGCGCTAGGCGGCTGCTTCAATTCAGTTTGAACCGCTGAGAGAC 660  
QY 600 TAAACGACGCGCGCTAGGCGGCTGCTTCAATTCAGTTTGAACCGCTGAGAGAC 659  
DB 661 GCAAGCTGCGGCTGCAATGTTTTCAGCGTGAAGAGAGAGAGAGAGAGAGAG 720  
QY 660 GCAAGCTGCGGCTGCAATGTTTTCAGCGTGAAGAGAGAGAGAGAGAGAGAG 719  
DB 721 GCTGCAAGACGCGCTAGATTAACCTTAAGAAATTAATCATTTGAGAGCTTAA 780  
QY 720 GCTGCAAGACGCGCTAGATTAACCTTAAGAAATTAATCATTTGAGAGCTTAA 779  
DB 781 AGATATCATGCTTAATTAATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840  
QY 780 AGATATCATGCTTAATTAATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 839  
DB 841 TGGGTACCGGCGCGCGCTGAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAG 900  
QY 840 TGGGTACCGGCGCGCGCTGAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAG 899

QY 901 AGCCCGGGGATCACTAGTCTAGAGCGCGCAACCGCGGTGAGCTCAGCTTTT 960  
DB 900 AGCCCGGGGATCACTAGTCTAGAGCGCGCGCAACCGCGGTGAGCTCAGCTTTT 959  
QY 961 TCCCTTATGAGAGGTTAATTAAGTCCGATGCGCAATTTTACGAGACATCTTTT 1020  
DB 960 TCCCTTATGAGAGGTTAATTAAGTCCGATGCGCAATTTTACGAGACATCTTTT 1019  
QY 1021 GATTATCTAGCTGATCAGATCATATCTGCGGCTTTTTCGCGCTCAGTATGCC 1080  
DB 1020 GATTATCTAGCTGATCAGATCATATCTGCGGCTTTTTCGCGCTCAGTATGCC 1079  
QY 1081 CAAGCTGCGCTATCTGGGATTCGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140  
DB 1080 CAAGCTGCGCTATCTGGGATTCGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1139  
QY 1141 AAGGCGATGGAAGAGTTGCGGAGATGACTGCTGCAATGAAGCTTGAAGCAAAAC 1200  
DB 1140 AAGGCGATGGAAGAGTTGCGGAGATGACTGCTGCAATGAAGCTTGAAGCAAAAC 1199  
QY 1201 GCAAGTTTACATGATGATTCGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260  
DB 1200 GCAAGTTTACATGATGATTCGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1259  
QY 1261 TCGTTCAAGGCGCACTGGGATACAGTTCGTGCGGCGCTTTTCGAGACAGTTCC 1320  
DB 1260 TCGTTCAAGGCGCACTGGGATACAGTTCGTGCGGCGCTTTTCGAGACAGTTCC 1319  
QY 1321 TCAGCCCGAGCGCATCAGACACCGGACAAATACCGGCGCAAGCGGAACTGCGG 1380  
DB 1320 TCAGCCCGAGCGCATCAGACACCGGACAAATACCGGCGCAAGCGGAACTGCGG 1379  
QY 1381 GTGTGCAATTAATGAAGCGGAGTGGGCTGGGATTAAGTCAAGAGAGAGAGATC 1440  
DB 1380 GTGTGCAATTAATGAAGCGGAGTGGGCTGGGATTAAGTCAAGAGAGAGAGATC 1439  
QY 1441 CTGGCTGATGCGCGAGAAATGAAGATGATACCGGCTGAGTTTCCGCGGCGGCGCT 1500  
DB 1440 CTGGCTGATGCGCGAGAAATGAAGATGATACCGGCTGAGTTTCCGCGGCGGCGCT 1499  
QY 1501 GGGGTAAATGATGATGATGCTGCTTCTGTGTAATTTGTTATCCGCTCAATTTCA 1560  
DB 1500 GGGGTAAATGATGATGATGCTGCTTCTGTGTAATTTGTTATCCGCTCAATTTCA 1559  
QY 1561 CAACATACGAGCGGAGAGATTAAGTGAAGCTGCGGAGCTTAAGAGAGAGAGAG 1620  
DB 1560 CAACATACGAGCGGAGAGATTAAGTGAAGCTGCGGAGCTTAAGAGAGAGAGAG 1619  
QY 1621 CAACTTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1679  
DB 1620 CAACTTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1679  
QY 1681 GCATTATGAATGAGCGCAAGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1740  
DB 1680 GCATTATGAATGAGCGCAAGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1739  
QY 1741 TTCTCGCTCACTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800  
DB 1740 TTCTCGCTCACTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1799  
QY 1801 CTCAAGGCGGATTAATGAGTTATCAAGATCAAGGAGATGAAGAGAGAGAGAGAGAG 1860  
DB 1800 CTCAAGGCGGATTAATGAGTTATCAAGATCAAGGAGATGAAGAGAGAGAGAGAGAG 1859  
QY 1861 AGCAAAAGCGCAAGAAAGCGCAAGAAAGCGCAAGAAAGCGCAAGAAAGCGCAAG 1920  
DB 1860 AGCAAAAGCGCAAGAAAGCGCAAGAAAGCGCAAGAAAGCGCAAGAAAGCGCAAG 1919  
QY 1921 TAGGCTCGGCGCGCTGAGAGATCAAAATTCAGCTCAAGTCAAGAGTGGGAGAA 1980  
DB 1920 TAGGCTCGGCGCGCTGAGAGATCAAAATTCAGCTCAAGTCAAGAGTGGGAGAA 1979

Qy	1981	CCGCAcAGsAaCTAAaAAGATACcAGGGGCTTCCCGCTGGAAGCTCCCTCGGCGCTCC	2040
Db	1980	CCCCAcAGsAGATTAaAAGATACcAGGGGTTCCCGCTGGAAGCTCCCTCGGCGCTCC	2039
Qy	2041	TGTTCCGAACCCCTGCGCGCTTACCGGATACCTGTCCGCTTCTCCCTTCGGGAACGTGGC	2100
Db	2040	TGTTCCGAACCCCTGCGCGCTTACCGGATACCTGTCCGCTTCTCCCTTCGGGAACGTGGC	2099
Qy	2101	GCCTTCTCATAGCTCACGCTGTAGGTATCTCAGTTGCGGTGAGGTGCTTCGAAGCT	2160
Db	2100	GCCTTCTCATAGCTCACGCTGTAGGTATCTCAGTTGCGGTGAGGTGCTTCGAAGCT	2159
Qy	2161	GGGCGTGTGTGACGAACCCCGCTTCAGCCCGACCCGCTGGCGCTTAACCGGTAACTATC	2220
Db	2160	GGGCGTGTGTGACGAACCCCGCTTCAGCCCGACCCGCTGGCGCTTAACCGGTAACTATC	2219
Qy	2221	TCTTGAGTCCAAACCCGGTAAAGACACGACTTATGCGCACTGGCAGCAGCACTGGTAACAG	2280
Db	2220	TCTTGAGTCCAAACCCGGTAAAGACACGACTTATGCGCACTGGCAGCAGCACTGGTAACAG	2279
Qy	2281	GATTAGCAGACGAGGTATGTAGCGCGGTCTACAGAGTTCTTGAAGTGTGGCTTAACTA	2340
Db	2280	GATTAGCAGACGAGGTATGTAGCGCGGTCTACAGAGTTCTTGAAGTGTGGCTTAACTA	2339
Qy	2341	CGGCTACACTAAGAAAGACAGTATTGGTATCTGCGCTCTGCTGAGCAGTAACTCTTCG	2400
Db	2340	CGGCTACACTAAGAAAGACAGTATTGGTATCTGCGCTCTGCTGAGCAGTAACTCTTCG	2399
Qy	2401	AAAAAGAGTGTGTAGCTCTTGAATCCGGCAAAACAACACGCGCTGTAGCGGTGTTTTT	2460
Db	2400	AAAAAGAGTGTGTAGCTCTTGAATCCGGCAAAACAACACGCGCTGTAGCGGTGTTTTT	2459
Qy	2461	TGTTTGCAGACAGCAGATTACGCCAGAAAAAAGATCTCTCAAGAAATCCTTTGATCTT	2520
Db	2460	TGTTTGCAGACAGCAGATTACGCCAGAAAAAAGATCTCTCAAGAAATCCTTTGATCTT	2519
Qy	2521	TTCTACGGGGCTCTGACGCTCAGTGAACGAAAACTCAGTTAAGGAGATTTTGTATGAG	2580
Db	2520	TTCTACGGGGCTCTGACGCTCAGTGAACGAAAACTCAGTTAAGGAGATTTTGTATGAG	2579
Qy	2581	ATTATCAAAAAAGATCTTCAACCTAGATCCTTTTAAATTTAAATGAAGCTTTTAAATCAAT	2640
Db	2580	ATTATCAAAAAAGATCTTCAACCTAGATCCTTTTAAATTTAAATGAAGCTTTTAAATCAAT	2639
Qy	2641	CTTAAAGTATATATAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGGAGCAC	2700
Db	2640	CTTAAAGTATATATAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGGAGCAC	2699
Qy	2701	TATCTCAGCGATCTGTCTATTTCTGTATCTCATCATATGTTGCTGACTCCCGCTCGTATGAT	2760
Db	2700	TATCTCAGCGATCTGTCTATTTCTGTATCTCATCATATGTTGCTGACTCCCGCTCGTATGAT	2759
Qy	2761	AACCTACGATACGGGAGGGCTTACCATCTGGGCCCAAGTCTGCAATGATACCGGAGACCC	2820
Db	2760	AACCTACGATACGGGAGGGCTTACCATCTGGGCCCAAGTCTGCAATGATACCGGAGACCC	2819
Qy	2821	ACGCTCACCGGCTCCAGATTAATACGCAATTAACACAGCAGCGGGAAGGGCGGCGCAG	2880
Db	2820	ACGCTCACCGGCTCCAGATTAATACGCAATTAACACAGCAGCGGGAAGGGCGGCGCAG	2879
Qy	2881	AAGTGTCTTGCAACTTATATCCGCTCATCTCAGTCTTAATATGTTTGGCGGGAAGCTAG	2940
Db	2880	AAGTGTCTTGCAACTTATATCCGCTCATCTCAGTCTTAATATGTTTGGCGGGAAGCTAG	2939
Qy	2941	AGTAAGTATGTTCCGCACTTAATAGTTTGGCGCAAGCTTGTTCATTCCTACAGGCATCT	3000
Db	2940	AGTAAGTATGTTCCGCACTTAATAGTTTGGCGCAAGCTTGTTCATTCCTACAGGCATCT	2999
Qy	3001	GGGTCAACGCTCGTGTGTGTGTAGGGCTTCAATCAGCTCCGGTCCCAAGATCAAGGCG	3060
Db	3000	GGGTCAACGCTCGT	3059
Qy	3061	AGTTACATGATCCCCACATGTTGTGCAAAAAAGCGGTTAGCTCTTCCGATCTCCGATCGT	3120

Db	3060	AGTTACATGATCCCCCATGTTTGCAAAAAGGGTTAGCTCCTTGGGTCTCCGATCGT	3119
Qy	3121	TGTCAGAAAGTAAGTTGGCCGACAGTGTATCATCTATGTTATGAGCACTGCATATATTC	3180
Db	3120	TGTCAGAAAGTAAGTTGGCCGACAGTGTATCATCTATGTTATGAGCACTGCATATATTC	3179
Qy	3181	TCTTACTGTCATGCGCATCCGTAAGATGCTTTTCTTGACTGTGTAGTACTCAACCAATGC	3240
Db	3180	TCTTACTGTCATGCGCATCCGTAAGATGCTTTTCTTGACTGTGTAGTACTCAACCAATGC	3239
Qy	3241	ATTCTGGAATATGATATGAGTATGGGGGACCGGATTCCTCTGCGCGCGGCGCATATAGGGATTA	3300
Db	3240	ATTCTGGAATATGATATGAGTATGGGGGACCGGATTCCTCTGCGCGCGGCGCATATAGGGATTA	3299
Qy	3301	TACCGCGCCCATATGACGAACTTTAAAGTGCTCATCATTTGGAAAAAGTTCTTTCGGGGCG	3360
Db	3300	TACCGCGCCCATATGACGAACTTTAAAGTGCTCATCATTTGGAAAAAGTTCTTTCGGGGCG	3359
Qy	3361	AAAACTCTCAAGATCTTTACCGCTGTTGAGATCCAGTTCGATGTAAACCACTGTCACCC	3420
Db	3360	AAAACTCTCAAGATCTTTACCGCTGTTGAGATCCAGTTCGATGTAAACCACTGTCACCC	3419
Qy	3421	CAACTGATCTTCAGCATCTTTTACTTTCACCAAGCTTTCTGGGTGAGCAAAAACAGGAAG	3480
Db	3420	CAACTGATCTTCAGCATCTTTTACTTTCACCAAGCTTTCTGGGTGAGCAAAAACAGGAAG	3479
Qy	3481	GCAAAATGCGGCAAAAAGGGAATTAAGGGGGAACGGAATGTTGAATACTCATACTCTT	3540
Db	3480	GCAAAATGCGGCAAAAAGGGAATTAAGGGGGAACGGAATGTTGAATACTCATACTCTT	3539
Qy	3541	CCTTTTCATATATATTTGAAGCATTTATCAAGGGTTATGTCTCATGAAGCGGATACATATT	3600
Db	3540	CCTTTTCATATATATTTGAAGCATTTATCAAGGGTTATGTCTCATGAAGCGGATACATATT	3599
Qy	3601	TGAATGTAATTTAGAAAAATTAACAATATGGGGTTCCGCGACATTTCCCGAAAAAGTGC	3660
Db	3600	TGAATGTAATTTAGAAAAATTAACAATATGGGGTTCCGCGACATTTCCCGAAAAAGTGC	3659
Qy	3661	AC 3662	
Db	3660	AC 3661	

RESULT 2

ADM68463/c

ID ADM68463 standard; DNA, 3637 BP.

XX ADM68463;

XX

DT 03-JUN-2004 (first entry)

XX

DE Variant green fluorescent protein, GFP, gene construct.

XX ds; mismatch endonuclease; endonuclease; gene shuffling technology;

KM single nucleotide polymorphism; cancer susceptibility;

XX sequence variation redistribution; GFP; green fluorescent protein; gene.

OS Aequorea victoria.

OS Synthetic.

XX

PN US2003157682-A1.

XX

PD 21-AUG-2003.

XX

PF 31-JAN-2003; 2003US-00356708.

XX

PR 01-FEB-2002; 2002US-0353722P.

PR 14-MAR-2002; 2002US-00098155.

PR 01-AUG-2002; 2002US-00211079.

XX

PA (PADG/) PADGETT H S.

PA (VAEW/) VALEWONGS A A.

PA (VOJ/D/) VOJDANI F S.  
PA (SMIT/) SMITH M L.  
PA (LIND/) LINDBO J A.  
PA (FITZ/) FITZMAURICE W P.  
XX  
XX  
PI Padgett HS, Vaeqhongs AA, Vojdani FS, Smith ML, Lindbo JA;  
PI Fitzmaurice WP;  
XX  
XX  
DR WPI, 2003-766176/72.  
XX  
XX  
PT Making a mismatch endonuclease, useful in gene shuffling and in detection  
PT of single nucleotide polymorphisms, comprises transfecting a host with a  
PT recombinant viral vector including a polynucleotide encoding a mismatch  
PT endonuclease.  
XX  
XX  
PS Example 5; SEQ ID NO 32; 79pp; English.  
XX

1620 TCACATTAATGCGTGGCTCACTGCCCCGCTTCCAGTCGGGAAACCTGTCGTCACG 1679  
1643 TCACATTAATGCGTGGCTCACTGCCCCGCTTCCAGTCGGGAAACCTGTCGTCACG 1684  
1680 TGCATTAATGAATCGCCCAACGCGCGGGAGAGCGGCTTTCGATTTGGGCGCTTTCG 1739  
1983 TGCATTAATGAATCGCCCAACGCGCGGGAGAGCGGCTTTCGATTTGGGCGCTTTCG 1924  
1740 CTTCCTCGCTCACTGACTCGCTGCGCTCGCTGCTTCGCTCGCTCGCGGAGCGGATCACTC 1799  
1923 CTTCCTCGCTCACTGACTCGCTGCGCTCGCTGCTTCGCTCGCTCGCGGAGCGGATCACTC 1864  
1800 ACTCAAGGCGGTAATACGCTTATCCAGAAATCAGAGGGATTAACGAGAAAGAACTAT 1859  
1863 ACTCAAGGCGGTAATACGCTTATCCAGAAATCAGAGGGATTAACGAGAAAGAACTAT 1804  
1860 GAGCAAAAGCGCAGCAAAAGCGCAGAAACCGTAAAGGCGCGCTTTCGCTGCTTTCG 1919  
1803 GAGCAAAAGCGCAGCAAAAGCGCAGAAACCGTAAAGGCGCGCTTTCGCTGCTTTCG 1744  
1920 ATAGGCTCCGCCCCCTGACGAGCATCAAAATCGACGCTCAAGTCAAGTGGCGAA 1979  
1743 ATAGGCTCCGCCCCCTGACGAGCATCAAAATCGACGCTCAAGTCAAGTGGCGAA 1684  
1980 ACCGACAGGACTATTAAGATACAGAGGCTTCCCGCTGGAAGCTCCCTCGTGGCTTC 2039  
1683 ACCGACAGGACTATTAAGATACAGAGGCTTCCCGCTGGAAGCTCCCTCGTGGCTTC 1624  
2040 CTGTTCCGACCCCTCGCTTACCGGATACGCTGCGCTTTCCTTCCTTCGGAAGCTGG 2059  
1623 CTGTTCCGACCCCTCGCTTACCGGATACGCTGCGCTTTCCTTCCTTCGGAAGCTGG 1564  
2100 CGCTTCTCATAGCTCACTGCTGATGATCTCAAGTTCGCTGATGAGCTGCTTCCCTCAAGC 2159  
1563 CGCTTCTCATAGCTCACTGCTGATGATCTCAAGTTCGCTGATGAGCTGCTTCCCTCAAGC 1504  
2160 TGGGCTGCTGCAAGAACCCCGCTTACGCGCGACCGCTGCGCTTATCCCGTAACTATC 2219  
1503 TGGGCTGCTGCAAGAACCCCGCTTACGCGCGACCGCTGCGCTTATCCCGTAACTATC 1444  
2220 GTCTTGAATCAACCCCGTAAAGACAGCACTTATCGCACTGAGCAGCAGCACTGGTAA 2279  
1443 GTCTTGAATCAACCCCGTAAAGACAGCACTTATCGCACTGAGCAGCAGCACTGGTAA 1384  
2280 GGAATTAAGCAGAGGATATGAGCGGCTGCTACAGAGTTCTGAAGTGGCTTAACT 2339  
1383 GGAATTAAGCAGAGGATATGAGCGGCTGCTACAGAGTTCTGAAGTGGCTTAACT 1324  
2340 ACCGCTACCTAAGAAAGCAATTTGGTATCTGCGCTTCGCTGAGCACTTACCTTCG 2399  
1323 ACCGCTACCTAAGAAAGCAATTTGGTATCTGCGCTTCGCTGAGCACTTACCTTCG 1264  
2400 GAAAGAGGTTGATGCTTGGATCCGCAAAACCAACCGCTGCTGAGCGGTTTTCG 2459  
1263 GAAAGAGGTTGATGCTTGGATCCGCAAAACCAACCGCTGCTGAGCGGTTTTCG 1204  
2460 TTGTTTGAAGCAGCAGATTAACGCGCAGAAAGAAAGATCTCAAGAGATCTTGAATCT 2519  
1203 TTGTTTGAAGCAGCAGATTAACGCGCAGAAAGAAAGATCTCAAGAGATCTTGAATCT 1144  
2520 TTCTTAAGGCTGAGCGCTCAAGGAAAGAAAGCTCACTTAAGGATTTTGGTCAATGA 2579  
1143 TTCTTAAGGCTGAGCGCTCAAGGAAAGAAAGCTCACTTAAGGATTTTGGTCAATGA 1084  
2580 GATTATCAAAAGGATCTTCACTAGATCTTTTAAATTAAGAAAGTTTAAATCA 2639  
1083 GATTATCAAAAGGATCTTCACTAGATCTTTTAAATTAAGAAAGTTTAAATCA 1024  
2640 TCTAAGTATATAGTAATACTTGGTCTGACAGTTACCAATGCTTAATCAGTAGGAC 2699  
1023 TCTAAGTATATAGTAATACTTGGTCTGACAGTTACCAATGCTTAATCAGTAGGAC 964  
2700 CTATCTCAGGATCTGCTATTTGGTTCATCAGTAGTTGCCGCTCCCGCTCGTGA 2759

963 CTATCTCAGGATCTGCTATTTTGGTTCATCAGTAGTTGCCGCTCCCGCTGTAAGA 904  
2760 TAACTAGATACGGAGGCGCTTACATCTGCCCCAGTGTGCAATAGATACCGGAGAC 2819  
903 TAACTAGATACGGAGGCGCTTACATCTGCCCCAGTGTGCAATAGATACCGGAGAC 844  
2820 CAGCTCAAGGCTCCAGATTTATCAGCAATTAACAGCCAGCGGAAAGGCGAGGCA 2879  
843 CAGCTCAAGGCTCCAGATTTATCAGCAATTAACAGCCAGCGGAAAGGCGAGGCA 784  
2880 GAAAGTGTCCGCAACTTTATCCGCTCAATCCAGTGTATTAATTTGTCGCGGAGCTA 2939  
783 GAAAGTGTCCGCAACTTTATCCGCTCAATCCAGTGTATTAATTTGTCGCGGAGCTA 724  
2940 GAGTAAGTATGTCAGGATTAATAGTTTGGCAACGTTGTCATTCATGCTACAGGATCG 2999  
723 GAGTAAGTATGTCAGGATTAATAGTTTGGCAACGTTGTCATTCATGCTACAGGATCG 664  
3000 TGGTGTCAAGCTGCTGCTTGGTATGCTTCAATCACTCGGATCCCAACGATCAAGGC 3059  
663 TGGTGTCAAGCTGCTGCTTGGTATGCTTCAATCACTCGGATCCCAACGATCAAGGC 604  
3060 GAGTTACATGATCCCGCATGTTGCAAAAGGCGTTAGGCTCCTTGGCTCCGATCG 3119  
603 GAGTTACATGATCCCGCATGTTGCAAAAGGCGTTAGGCTCCTTGGCTCCGATCG 544  
3120 TTGTCAAGATGATGTCGCGCAGTGTATCACTCATGATGATTAAGCAGCACTGATTAAT 3179  
543 TTGTCAAGATGATGTCGCGCAGTGTATCACTCATGATGATTAAGCAGCACTGATTAAT 484  
3180 CTCTTATCTGATGCTCATCCGTAAGATGCTTTTCTGATCTGATGATGATCAACAACT 3239  
483 CTCTTATCTGATGCTCATCCGTAAGATGCTTTTCTGATCTGATGATGATCAACAACT 424  
3240 CATTCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3299  
423 CATTCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 364  
3300 ATACCGCGCAGATAGCAGAACTTTAAAGTGTCACTCATTTGAAACGTTCTTCGCGGC 3359  
363 ATACCGCGCAGATAGCAGAACTTTAAAGTGTCACTCATTTGAAACGTTCTTCGCGGC 304  
3360 GAAAGCTCAAGGATCTTACCGCTGTTGATGATCACTGATGATCAACCACTGTCAC 3419  
303 GAAAGCTCAAGGATCTTACCGCTGTTGATGATCACTGATGATCAACCACTGTCAC 244  
3420 CCAACTGATCTTCAAGATCTTTTCACTTCAACGAGCTTCTGCGGTGAGCAAAACAGAA 3479  
243 CCAACTGATCTTCAAGATCTTTTCACTTCAACGAGCTTCTGCGGTGAGCAAAACAGAA 184  
3480 GGCAGAAATGCGCAAAAGGAAATAGGCGCAGCAACGGAATGTTGAATCTCACTCT 3539  
183 GGCAGAAATGCGCAAAAGGAAATAGGCGCAGCAACGGAATGTTGAATCTCACTCT 124  
3540 TCCCTTTTCAATATTAAGCAATTAATCAAGGTTATTTGTCTCATAGGCGATCAATAT 3599  
123 TCCCTTTTCAATATTAAGCAATTAATCAAGGTTATTTGTCTCATAGGCGATCAATAT 64  
3600 TTGAATGATTTGAAATTAACAAATAGGCGGTCCGCGCACTTCCCGGAAAGATGC 3659  
63 TTGAATGATTTGAAATTAACAAATAGGCGGTCCGCGCACTTCCCGGAAAGATGC 4  
3660 CAC 3662  
3 CAC 1  
RESULT 3  
ADP26616/c  
ID ADP26616 standard; DNA; 3637 BP.  
XX  
AC ADP26616;

Oy

5

Ph

2

Dih

Q

Di

QY

10

7

79

9b

4

0

7

1



```
Db 2382 TGGCCATGGAAACAGGTATGTTTCCAGTAGTGTGCAAAATTAATTTAAGTGAAGCTTTCGGTA 2322
Oy 1330 GTCAAGCCCGAAGCGCATGACAAACCGAAATACCGCGGACAGCGCGGAACCTGCGTGC 1379
Db 2332 TGATGATCACCCTTCAACCTCTCCACTGACAGAAAATTTGTGCCATTAAATCATCAACATC 2263
Oy 1380 GGTGTGAGATTAAAGACAGCGGTGGCGCTGGGATTTTACGTACAGGAGACGAGGTAT 1439
Db 2262 TAAATTCAACAAGAAAT-----TGGGACAACTTCAGTGAAGAAAGTTC 2224
Oy 1440 CCTGGCTGGATGGCGCAGAAAATGACATGATACCCCGGTGAGTTACCCGGCGGGCGGCT 1499
Db 2223 TTCTCCTTACTCATCGGTATCCAGCTTTTGTTCCTTTAGTGAAGGTTAATTCGCGCT 2164
Oy 1500 TGGCGTAATCATGCTATAGCTGTTTCTGTGTGTAATTTGTTATCCGTCACAAATTCAC 1559
Db 2163 TGGCGTAATCATGCTATAGCTGTTTCTGTGTGTAATTTGTTATCCGTCACAAATTCAC 2104
Oy 1560 ACAACATACGAGCCGGAAGCATTAAGTGTAAAGCTGGGGGTGCTTAATGATGAGCTAAC 1619
Db 2103 ACAACATACGAGCCGGAAGCATTAAGTGTAAAGCTGGGGGTGCTTAATGATGAGCTAAC 2044
Oy 1620 TCACATTAAATTGCGTTGCGTCACTGCGCGCTTTCCAGTCGCGGAAAACCTGTCGTCAGC 1679
Db 2043 TCACATTAAATTGCGTTGCGTCACTGCGCGCTTTCCAGTCGCGGAAAACCTGTCGTCAGC 1984
Oy 1680 TGCATTAAATGAAATGCGCCAAACGCGCGGGGAGAGCGGCTTTGCGTATTTGGGCGCTTC 1739
Db 1983 TGCATTAAATGAAATGCGCCAAACGCGCGGGGAGAGCGGCTTTGCGTATTTGGGCGCTTC 1924
Oy 1740 CTTCCTGCTCATGCTTCGCTGCGCTGCGCTGCGCTGCGCTGCGGAGACGGTATACGTC 1799
Db 1923 CTTCCTGCTCATGCTTCGCTGCGCTGCGCTGCGCTGCGCTGCGGAGACGGTATACGTC 1864
Oy 1800 ACTCAAAAGCGGTAATACGTTTATCACAAGATCAGGGGATTAACAGAGAAAGAACATGT 1859
Db 1863 ACTCAAAAGCGGTAATACGTTTATCACAAGATCAGGGGATTAACAGAGAAAGAACATGT 1804
Oy 1860 GAGCAAAAAGCCAGCAAAAAGCCAGGAACCGTAAAGAGCGCGCTGCTGCGCTTTTC 1919
Db 1803 GAGCAAAAAGCCAGCAAAAAGCCAGGAACCGTAAAGAGCGCGCTGCTGCGCTTTTC 1744
Oy 1920 ATAGGCTCCGCCCCCTTACAGAGATCACAAGAAATCGACGCTCAAGTCAAGGTGCGAA 1979
Db 1743 ATAGGCTCCGCCCCCTTACAGAGATCACAAGAAATCGACGCTCAAGTCAAGGTGCGAA 1684
Oy 1980 ACCGACAGAGACTATTAAGATACAGAGGCTTTCCCGTGAAGGCTCCCTGCGGCTTC 2039
Db 1683 ACCGACAGAGACTATTAAGATACAGAGGCTTTCCCGTGAAGGCTCCCTGCGGCTTC 1624
Oy 2040 CTGTTCCGACCCCTGCGCTTACCGGATACCTGTGCGCTTTCTCCCTTCGGAAGCGTGG 2099
Db 1623 CTGTTCCGACCCCTGCGCTTACCGGATACCTGTGCGCTTTCTCCCTTCGGAAGCGTGG 1564
Oy 2100 CGCTTTTCATAGCTCAGGCTGTAGGTATCTCAAGTTGCGGTAGGTGCTTCCCTCAAGC 2159
Db 1563 CGCTTTTCATAGCTCAGGCTGTAGGTATCTCAAGTTGCGGTAGGTGCTTCCCTCAAGC 1504
Oy 2160 TGGGCTGTGTGCAAGAACCCCGCTTACGCGGACCGGCTGCGCTTATCGGTAATATC 2219
Db 1503 TGGGCTGTGTGCAAGAACCCCGCTTACGCGGACCGGCTGCGCTTATCGGTAATATC 1444
Oy 2220 GTCTTGAATCAACCCCGGTAAAGACAGCTTATCGCACTGGCAGACAGCACTGGTATCA 2279
Db 1443 GTCTTGAATCAACCCCGGTAAAGACAGCTTATCGCACTGGCAGACAGCACTGGTATCA 1384
Oy 2280 GGAATTAAGAGCGAGGTATGTAGCGGTGCTACAGAGTTCTTGAAGGTGCGCTTAAT 2339
Db 1383 GGAATTAAGAGCGAGGTATGTAGCGGTGCTACAGAGTTCTTGAAGGTGCGCTTAAT 1324
Oy 2340 ACGGCTACATAGAGAGCAGTATTTGGTATCTGCGCTCTGCTGAAGCAGTTACTCTCG 2399
Db 1323 ACGGCTACATAGAGAGCAGTATTTGGTATCTGCGCTCTGCTGAAGCAGTTACTCTCG 1264
Oy 2400 GAAAAAGAGTTGGTAGCTTTGATCCGCAAAACCAACCGCTGGTAGCGGTGTTTTT 2459
Db 1263 GAAAAAGAGTTGGTAGCTTTGATCCGCAAAACCAACCGCTGGTAGCGGTGTTTTT 1204
Oy 2460 TTGTTTCAACACACAGATTACCGGCGCAAAAAAAGATCTCAAGAAATCCCTTGATCT 2519
Db 1203 TTGTTTCAACACAGATTACCGGCGCAAAAAAAGATCTCAAGAAATCCCTTGATCT 1144
Oy 2520 TTTCTACGGGGTCTGACGCTCAGTGAAGCAAAAACTCAGCTTAAGGATTTTGGTCA 2579
Db 1143 TTTCTACGGGGTCTGACGCTCAGTGAAGCAAAAACTCAGCTTAAGGATTTTGGTCA 1084
Oy 2580 GATTATCAAAAAGGATCTTCACTAGATCCCTTTAAATTAATAATGAAGTTTAAATCA 2639
Db 1083 GATTATCAAAAAGGATCTTCACTAGATCCCTTTAAATTAATAATGAAGTTTAAATCA 1024
Oy 2640 TCTAAGATATTAAGATTAACCTTGTCTGACAGTTACCAATGCTTAATCAAGAGCAC 2699
Db 1023 TCTAAGATATTAAGATTAACCTTGTGTGACAGTTACCAATGCTTAATCAAGAGCAC 964
Oy 2700 CTATCTCAGGATCTGTCTATTTGCTCATTCATAGTTGCTGACTCCCGCTGCTAGA 2759
Db 963 CTATCTCAGGATCTGTCTATTTGCTCATTCATAGTTGCTGACTCCCGCTGCTAGA 904
Oy 2760 TAACTACGATACGGAGAGGCTTACCATCTGCGCCAGTGTGCAATGATACCGGAGACC 2819
Db 903 TAACTACGATACGGAGAGGCTTACCATCTGCGCCAGTGTGCAATGATACCGGAGACC 844
Oy 2820 CACGCTCACCGGCTCCAGATTTATACAGCAATTAACAGCCAGCGGAAAGGCGGACGCA 2879
Db 843 CACGCTCACCGGCTCCAGATTTATACAGCAATTAACAGCCAGCGGAAAGGCGGACGCA 784
Oy 2880 GAAATGCTCTGCAACTTTATCCGCTCCATCCAGTCTATTAATTTGTGCGGGAAGCTA 2939
Db 783 GAAATGCTCTGCAACTTTATCCGCTCCATCCAGTCTATTAATTTGTGCGGGAAGCTA 724
Oy 2940 GAGTAAGTATGTCGCAAGTTAATAGTTTGGCGCAAGTTTGGCCATTTGCTACAGGATCG 2999
Db 723 GAGTAAGTATGTCGCAAGTTAATAGTTTGGCGCAAGTTTGGCCATTTGCTACAGGATCG 664
Oy 3000 TGTGTCAAGCTGCTGCTTTGTGTATGCTTCACTTCACTCCGCTCCCAACGATCAAGGC 3059
Db 663 TGTGTCAAGCTGCTGCTTTGTGTATGCTTCACTTCACTCCGCTCCCAACGATCAAGGC 604
Oy 3060 GAGTTACATGATCCCATGTTGTCGCAAAAAGGAGGTAGTCTTCCGCTCCGCTCGATCG 3119
Db 603 GAGTTACATGATCCCATGTTGTCGCAAAAAGGAGGTAGTCTTCCGCTCCGCTCGATCG 544
Oy 3120 TTGTCAAGATTAAGTTGGCGCAGTGTATCACTCATGTTATGGCAGCATGCTGATAT 3179
Db 543 TTGTCAAGATTAAGTTGGCGCAGTGTATCACTCATGTTATGGCAGCATGCTGATAT 484
Oy 3180 CTCTTACTGTATGCTCATCCGTAAGATGCTTTTCTGTGACTGTGAGTATCAACCAAGT 3239
Db 483 CTCTTACTGTATGCTCATCCGTAAGATGCTTTTCTGTGACTGTGAGTATCAACCAAGT 424
Oy 3240 CATTCAGAAATAGTATGTCGCGGACAGAGTTGCTTTCGCGCGGCTCAATACGGGATA 3299
Db 423 CATTCAGAAATAGTATGTCGCGGACAGAGTTGCTTTCGCGCGGCTCAATACGGGATA 364
Oy 3300 ATACCGGCGCAATAGCAGAACTTTAAAGTGTCTCATCTTGAAGAAAGCTTTCGCGGC 3359
Db 363 ATACCGGCGCAATAGCAGAACTTTAAAGTGTCTCATCTTGAAGAAAGCTTTCGCGGC 304
Oy 3360 GAAAACTCTCAAGATCTTACCGCTGTGAGATCCAGTTGAGATGAACCACTCTGTAC 3419
Db 303 GAAAACTCTCAAGATCTTACCGCTGTGAGATCCAGTTGAGATGAACCACTCTGTAC 244
Oy 3420 CCAACTGATCTTCAGCATCTTTTACTTCAACAGCGTTTCTGGGTGAGCAAAAACAGAA 3479
Db 243 CCAACTGATCTTCAGCATCTTTTACTTCAACAGCGTTTCTGGGTGAGCAAAAACAGAA 184
```

Qy 3480 GGCATAATGCCGCAAAAAGGGAATTAAGGCGACACGAAATGTGAATACATCTCT 3539  
Db 183 GGCATAATGCCGCAAAAAGGGAATTAAGGCGACACGAAATGTGAATACATCTCT 124  
Qy 3540 TCCTTTTTCATATATTTGAAGCATTTATCAGGTTATTTCTCATGAGCGGATACATAT 3599  
Db 123 TCCTTTTTCATATATTTGAAGCATTTATCAGGTTATTTCTCATGAGCGGATACATAT 64  
Qy 3600 TTGAATGATTTAGAAAAATTAACAATAGGGGTTCCGCGACATTTCCCGAAAAAGTGC 3659  
Db 63 TTGAATGATTTAGAAAAATTAACAATAGGGGTTCCGCGACATTTCCCGAAAAAGTGC 4  
Qy 3660 CAC 3662  
Db 3 CAC 1

RESULT 4  
ABX14571/c  
ID ABX14571 standard; DNA; 3637 BP.  
AC ABX14571;  
XX  
XX 14-MAR-2003 (first entry)  
XX  
XX Plasmid pBSC3GFP DNA.  
XX  
XX Circular; cyclic; heteroduplex; T4 DNA polymerase; T4 DNA ligase; CEL I;  
XX population diversity; exonuclease; complementarity; de.  
XX Synthetic.  
XX WO200279468-A2.  
XX  
XX 10-OCT-2002.  
XX  
XX 01-FEB-2002; 2002MO-US03055.  
XX  
XX 02-FEB-2001; 2001US-0266386P.  
XX 14-FEB-2001; 2001US-0268785P.  
XX  
XX (LARG-) LARGE SCALE BIOLOGY CORP.  
XX  
XX Padgett HS, Fitzmaurice WP, Lindo JA;  
XX WPI; 2003-092898/08.  
XX  
XX Increasing complementarity by making sequence variants from heteroduplex  
XX polynucleotides, useful for producing improved gene products from  
XX randomly mutated genes.  
XX  
XX Example 5; Page 99-100; 105pp; English.  
XX  
XX The invention relates to an in vitro method of making sequence variants  
XX from at least one heteroduplex polynucleotide that has at least two-non  
XX complementary nucleotide base pairs. The method comprises combining a  
XX heteroduplex polynucleotide with CEL I, T4 DNA polymerase and T4 DNA  
XX ligase, or an agent or agents with exonuclease activity and allowing  
XX sufficient time for the percentage of complementarity to increase, where  
XX one or more variants are made. The method can also be used to increase  
XX diversity in a population of sequences. The method is useful for  
XX expressing improved gene products from randomly mutated genes or from  
XX expression from a suitable plant, animal, fungal, yeast or bacterial  
XX recombinant vector, and provides a high-efficiency recovery of  
XX CC that are more diverse and with a lower percentage of sequence identity.  
XX This sequence represents plasmid DNA, used in heteroduplex substrate  
XX preparation  
XX  
XX

Sequence 3637 BP; 942 A; 892 C; 870 G; 933 T; 0 U; 0 Other;  
Query Match 72.8%; Score 2664.8; DB 8; Length 3637;

Best Local Similarity 83.8%; Pred. No. 0;  
Matches 3070; Conservative 0; Mismatches 567; Indels 25; Gaps 4;

Qy 1 CTAAATTTAGACCTTAATATTTTGTAAATATTCGGTTAAATTTTGTAAATAGAGCTC 60  
Db 3637 CTAAATTTAGACCTTAATATTTTGTAAATATTCGGTTAAATTTTGTAAATAGAGCTC 3578  
Qy 61 ATTTTAAACCAATAGGCGCAAAATCGCAAAATCCCTTAATTAATCAAAAGATAGACCGA 120  
Db 3577 ATTTTAAACCAATAGGCGCAAAATCGCAAAATCCCTTAATTAATCAAAAGATAGACCGA 3518  
Qy 121 GATAGGTTGATGTTGTTCCAGTTTGAACAAGAGTCCACTATTAAAGAAAGTGAAGTCC 180  
Db 3517 GATAGGTTGATGTTGTTCCAGTTTGAACAAGAGTCCACTATTAAAGAAAGTGAAGTCC 3458  
Qy 181 CAAGTCGCAAAAGGCGCAAAACCGTCTATCAGGCGATGCGCACCTAGTGAACATCACC 240  
Db 3457 CAAGTCGCAAAAGGCGCAAAACCGTCTATCAGGCGATGCGCACCTAGTGAACATCACC 3398  
Qy 241 CTATCAAGTTTATTTGGGTCGAGGTCGCTTAAGCACTAATGTGAACCTTAAGGAGAG 300  
Db 3397 CTATCAAGTTTATTTGGGTCGAGGTCGCTTAAGCACTAATGTGAACCTTAAGGAGAG 3338  
Qy 301 CCCCAGATTTAGAGCTTGAACGGGCAAAAGCGGCGAAGCTGCGCAAGAAAGGAAAGAA 360  
Db 3337 CCCCAGATTTAGAGCTTGAACGGGCAAAAGCGGCGAAGCTGCGCAAGAAAGGAAAGAA 3278  
Qy 361 AGCGAAGAGACGGGCGCTGAGGCGCTGCAAGGTGAAGCGTACCGTCCGCTAACAC 420  
Db 3277 AGCGAAGAGACGGGCGCTGAGGCGCTGCAAGGTGAAGCGTACCGTCCGCTAACAC 3218  
Qy 421 CACACCGCGCGGCTTAAATGCGCCGCTACAGGCGGCGTCCATTCAGCTCAGCTGCG 480  
Db 3217 CACACCGCGCGGCTTAAATGCGCCGCTACAGGCGGCGTCCATTCAGCTCAGCTGCG 3158  
Qy 481 CAATCTTTGGAAGAGGCGATCGGTGCGGCTCTTTCGCTATTACGCCAGCTGCGGAAGG 540  
Db 3157 CAATCTTTGGAAGAGGCGATCGGTGCGGCTCTTTCGCTATTACGCCAGCTGCGGAAGG 3098  
Qy 541 GGGATGTGCTGCAAGGCGATTAAGTGGGTAAAGCGCAGGCTTTTCCAGTACGACGTTG 600  
Db 3097 GGGATGTGCTGCAAGGCGATTAAGTGGGTAAAGCGCAGGCTTTTCCAGTACGACGTTG 3038  
Qy 601 TAAACGACGCGCAAGTGAAGCGGCTCTTTCATTCAGCTTTTGAACCCGTGAGAGAG 660  
Db 3037 TAAACGACGCGCAAGTGAAGCGGCTCTTTCATTCAGCTTTTGAACCCGTGAGAGAG 2978  
Qy 661 GCAGACTCGGCTGCAAAATGTTTACAGGTGATGAGCAGATGAAGATGCTCGACAC 720  
Db 2977 CCGCGGTGGCGG-----CCGCTCTGAACCTAGTGAATCCCG--CGG 2939  
Qy 721 GCTGAGAAACACGACGCTAGATTAACCTTGAAGAAAGATATCATATGTAGACGATTGA 780  
Db 2938 GCTGAGAAATCTTATTTATTTATTTATTCATTCATTCATTCATTCATTCATTCATTCAT 2879  
Qy 781 AGATTAATCATCGTAATATTTGAAGCAATGGATCTTGAATACGATCTCATATGAGCGAT 840  
Db 2878 CAATCTCAAGAAAGCAATGTGTCACGCTTTCGTTGGATCTTTCGAAGGCGAGATT 2819  
Qy 841 TGGGTACCGGCGGCGGCTGAGGTGAGCGATGATGATGATGATGATGATGATGATGATG 900  
Db 2818 GTGTCGACAGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 900  
Qy 901 AGCCCGGGGATCAGTATGTTCTAGAGGCGCGCCACCGCGGTGAAGTCCAGCTTTTGT 2759  
Db 2758 GTTGAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 960  
Qy 961 TCCCTTGAAGAGGTTAATTAATCCATGCTCAATTTTACGACAGATCATCTTTAG 1020  
Db 2698 TACCTTTGAATTCATCTTTTGTGTTGTCGCGGATGATGATGATGATGATGATGATGATG 2640  
Qy 1021 GGTAAATCAAGCTGATCAGATCATATCGTGGGCTTTTTCGCGCTCAGTATCGCC 1080  
Db 1021 GGTAAATCAAGCTGATCAGATCATATCGTGGGCTTTTTCGCGCTCAGTATCGCC 1080

Db 2639 TTGATCTGAGTTGTGTCCGAGATGTTTCATCTCTTTAAATCAATACCTTTTAC 2580  
Qy 1081 CAAGTCGCGCTATCTGTGGCAATCGGGAGAGAAAGCCGTGCTTTTCCGCGAGTTG 1140  
Db 2579 TCGATACGATTAACAAGGGTATCACTTCAAACTTGACCTTCAGCAGCGCTCTTGTAGTTTC 2520  
Qy 1141 AAGCGGATGGAAGAGTTTGGCCGAGATGACTGCTGTCATTCAGCTTGAAGCGAANAAC 1200  
Db 2519 CCGTCATCTTTGAAAGATATAGTGTCTTCGTATATACCTTCGGGCAATGGCACTTTG 2460  
Qy 1201 GCACGTTTACATGATGATTCGGGAAGGTGTGGCCATCGCCCTTTATACGGTGAATGT 1260  
Db 2459 AAAAGTCATGCGGTTTATATATGATTCGGGATACGGGAAAAGCANTGAACACATAAAG 2400  
Qy 1261 TCGTTCAAGCCACCTGGGATACCACTTGTGCGGCTTTTCCGACACAGTTCCGATGG 1320  
Db 2399 AAAGTAGACAAAGTGTGGCCATGGAAACAGGTATTTTCCAGTAGTCCAAATAATTTA 2340  
Qy 1321 TCAGCCGGAAGGCAATCAGCAACCGGAACAATACGGGAGACGCCGGAACGTGCGG 1380  
Db 2339 AGGTAAGCTTTCCGTATGTAGCATCACCTTCACTTCCTCACTGACAGAAATTTTGC 2280  
Qy 1381 GTGTGCAATTAATGACAGCGGTGCGCGCTGGGATATTACGTACAGCAGACGGGTATC 1440  
Db 2279 CCATTAACATCACCATCTAATTCAACAAGAAATGGGACAATCCAGTGAAAAGTTCTTCT 2220  
Qy 1441 CTGGCTGATGCCGAGAAATGACATGAAATCCCGTAGTTACCCGCGGCGCGCTT 1500  
Db 2219 C---CTTACTCATCGGTACCCAGCTTTGTTCCTTTAGTAGGGTTAATTTGCGCGCTT 2163  
Qy 1501 GGGGTAATCATGTCATGCTGTTTCTGTGGAATTTGTTATCCGCTCACAAATTTCCCA 1560  
Db 2162 GGGGTAATCATGTCATGCTGTTTCTGTGGAATTTGTTATCCGCTCACAAATTTCCCA 2103  
Qy 1561 CAACATACGAGCCGGAAACATTAAGTGTAAAGCCTGGGCTCTAATGATGAGTAACT 1620  
Db 2102 CAACATACGAGCCGGAAACATTAAGTGTAAAGCCTGGGCTCTAATGATGAGTAACT 2043  
Qy 1621 CACATTAATTCGCTTGGCTCACTGCTCCGCTTTTCCAGTCGAGAAAACCTGTGTGTCAGCT 1680  
Db 2042 CACATTAATTCGCTTGGCTCACTGCTCCGCTTTTCCAGTCGAGAAAACCTGTGTGTCAGCT 1983  
Qy 1681 GCATTAATGAATCGGCAACGCGGGGAGAGCGGCTTTGGGTATGGGCGCTCTTCCG 1740  
Db 1982 GCATTAATGAATCGGCAACGCGGGGAGAGCGGCTTTGGGTATGGGCGCTCTTCCG 1923  
Qy 1741 TTCTCTGCTCACTGACTGCTGCGCTCGGTCTGTGCGCTGCGGAGCGGATACGCTCA 1800  
Db 1922 TTCTCTGCTCACTGACTGCTGCGCTCGGTCTGTGCGCTGCGGAGCGGATACGCTCA 1863  
Qy 1801 CTCAAAAGCGGTAATACGGTTATTCACAGAAATCAGGGGATACGAGGAAAAGATGTG 1860  
Db 1862 CTCAAAAGCGGTAATACGGTTATTCACAGAAATCAGGGGATACGAGGAAAAGATGTG 1803  
Qy 1861 AGCAAAAGGCGAGAAAAGGCGAGAAACCGTAAAGGCGCGCTTGGCTGGGCTTTTTC 1920  
Db 1802 AGCAAAAGGCGAGAAAAGGCGAGAAACCGTAAAGGCGCGCTTGGCTGGGCTTTTTC 1743  
Qy 1921 TAGGCTCGCGCCCTGACGAGCATCACAAAATTCAGCTCAAGTCAGAGGTGGCGAAA 1980  
Db 1742 TAGGCTCGCGCCCTGACGAGCATCACAAAATTCAGCTCAAGTCAGAGGTGGCGAAA 1683  
Qy 1981 CCGCAGAGGATTAAGATACAGAGCGTTTCCCGCTGGAAGCTCCCTGTGCGCTTCC 2040  
Db 1682 CCGCAGAGGATTAAGATACAGAGCGTTTCCCGCTGGAAGCTCCCTGTGCGCTTCC 1623  
Qy 2041 TGTTCGAGCCCTGCGCTTACCGGATACGTCGCGCTTTCCTTCCTTCGGGAAAGCTGGC 2100  
Db 1622 TGTTCGAGCCCTGCGCTTACCGGATACGTCGCGCTTTCCTTCCTTCGGGAAAGCTGGC 1563  
Qy 2101 GCTTCTCATAGCTACGCTGTAGGTATCTCAGTTCGGGTAGGTGCTTCCGCTCCAGCT 2160  
Db 1562 GCTTCTCATAGCTACGCTGTAGGTATCTCAGTTCGGGTAGGTGCTTCCGCTCCAGCT 1503

Qy 2161 GGGCTGTGTGACGAAACCCCGCTTACGCCGACCGCTGCGCTTATCCGGTAATCTATCG 2220  
Db 1502 GGGCTGTGTGACGAAACCCCGCTTACGCCGACCGCTGCGCTTATCCGGTAATCTATCG 1443  
Qy 2221 TCTTAGTTCACACCGGATTAAGACAGACTTATTCGCACTGGCAGCAGCACTGGTAAACAG 2280  
Db 1442 TCTTAGTTCACACCGGATTAAGACAGACTTATTCGCACTGGCAGCAGCACTGGTAAACAG 1383  
Qy 2281 GATTAGCAGAGGATATGTAGCGGTGTACAGAGTCTTGAAGTGTGGCTTAACCTA 2340  
Db 1382 GATTAGCAGAGGATATGTAGCGGTGTGTACAGAGTCTTGAAGTGTGGCTTAACCTA 1323  
Qy 2341 CCGCTACACTTAAGAGACAGTATTTGATATGTGGCTCTGCTGAAGCAGTACTTCTGG 2400  
Db 1322 CCGCTACACTTAAGAGACAGTATTTGATATGTGGCTCTGCTGAAGCAGTACTTCTGG 1263  
Qy 2401 AAAAAGATTGTAGCTTGTATCCGGCAAAACAACACCGCTGTAGCGGTGTTTTT 2460  
Db 1262 AAAAAGATTGTAGCTTGTATCCGGCAAAACAACACCGCTGTAGCGGTGTTTTT 1203  
Qy 2461 TGTTCGAGCAGCAGATTAACGCGCAGAAAAAAGATCTCAAGAGATCTTTGATCTT 2520  
Db 1202 TGTTCGAGCAGCAGATTAACGCGCAGAAAAAAGATCTCAAGAGATCTTTGATCTT 1143  
Qy 2521 TTCTACGGGCTCTGACGCTCAGTGAACGAAATCTACGTTAAGGATTTGTCTATGAG 2580  
Db 1142 TTCTACGGGCTCTGACGCTCAGTGAACGAAATCTACGTTAAGGATTTGTCTATGAG 1083  
Qy 2581 ATTATCAAAAAAGATCTTCACTGATCTTTTAAATTTAAAAAGAGTTTAAATCAAT 2640  
Db 1082 ATTATCAAAAAAGATCTTCACTGATCTTTTAAATTTAAAAAGAGTTTAAATCAAT 1023  
Qy 2641 CTAAAGTATATATAGTAAACCTTGTGTGACAGTTACCAATGCTTAAATCAAGTAGGCAAC 2700  
Db 1022 CTAAAGTATATATAGTAAACCTTGTGTGACAGTTACCAATGCTTAAATCAAGTAGGCAAC 963  
Qy 2701 TATCTCAGCATCTGTCTATTTGCTTATCCATATGTTGCTGACTCCCGCTGTGTAGAT 2760  
Db 962 TATCTCAGCATCTGTCTATTTGCTTATCCATATGTTGCTGACTCCCGCTGTGTAGAT 903  
Qy 2761 AACTACGATACGGGAGGGCTTACCATCGGCGCCAGTGTGCAATGATACCGGAGACC 2820  
Db 902 AACTACGATACGGGAGGGCTTACCATCGGCGCCAGTGTGCAATGATACCGGAGACC 843  
Qy 2821 ACGCTCACCGGCTTCAGATTTATCAGCAATAAACAGCACGCGGAAAGGCGGAG 2880  
Db 842 ACGCTCACCGGCTTCAGATTTATCAGCAATAAACAGCACGCGGAAAGGCGGAG 783  
Qy 2881 AAGTGTCTGCGAATTTATCCGCTTCATCCAGTCTAATTAATTTGTCGGGAAAGCTAG 2940  
Db 782 AAGTGTCTGCGAATTTATCCGCTTCATCCAGTCTAATTAATTTGTCGGGAAAGCTAG 723  
Qy 2941 AGTAAGTATTCGCACTTAATAGTTTGGCCAAAGTGTGGCTATTCCTACAGCATCTG 3000  
Db 722 AGTAAGTATTCGCACTTAATAGTTTGGCCAAAGTGTGGCTATTCCTACAGCATCTG 663  
Qy 3001 GGTGTCAAGCTCGTGTGGTATGAGCTTCAATTCAGCTCGGTTCCCAAGATCAAGGCG 3060  
Db 662 GGTGTCAAGCTCGTGTGGTATGAGCTTCAATTCAGCTCGGTTCCCAAGATCAAGGCG 603  
Qy 3061 AGTTACATGATCCCGCATGTTGTGCAAAAAAGCGGTTAGCTCTTGGGTCTCCGATCTG 3120  
Db 602 AGTTACATGATCCCGCATGTTGTGCAAAAAAGCGGTTAGCTCTTGGGTCTCCGATCTG 543  
Qy 3121 TGTGAGAGTATAGTTGGCCGAGTGTATACCTCATGCTATGAGCAGCAGCTAATATTC 3180  
Db 542 TGTGAGAGTATAGTTGGCCGAGTGTATACCTCATGCTATGAGCAGCAGCTAATATTC 483  
Qy 3181 TCTTACTGTATGCAATCCGTAAGATCTTTCTGTGATCTGTGAGTGAATCAACCAAGTC 3240  
Db 482 TCTTACTGTATGCAATCCGTAAGATCTTTCTGTGATCTGTGAGTGAATCAACCAAGTC 423

```

OY 3341 ATTCTGAGAAATAGTATGTCGGCGACCGAGTTGCTTTCCTCCCGGGGTCAATACGGGATTA 3300
Db 422 ATTCTGAGAAATAGTATGTCGGCGACCGAGTTGCTTTCCTCCCGGGGTCAATACGGGATTA 363
OY 3301 TACCGGGCCACATAGCAGAACTTTAAAGTGCATCATTTGAAAAAGTTCTTCGGGGCG 3360
Db 362 TACCGGGCCACATAGCAGAACTTTAAAGTGCATCATTTGAAAAAGTTCTTCGGGGCG 303
OY 3361 AAAACTCTGAGAGATTTTCCGCTGTGAGATCGAGTTGATGTAACCACTCGTGACC 3420
Db 302 AAAACTCTGAGAGATTTTCCGCTGTGAGATCGAGTTGATGTAACCACTCGTGACC 243
OY 3421 CAACGATTTTACGATCTTTTACTTTCACAGGGTTTCGGGTGACGAAAAACAGAGAG 3480
Db 242 CAACGATTTTACGATCTTTTACTTTCACAGGGTTTCGGGTGACGAAAAACAGAGAG 183
OY 3481 GCAAAATGCGCGAAAAAGGAAATAGGCGACACGGAATGTTGAATCTCATCTCTT 3540
Db 182 GCAAAATGCGCGAAAAAGGAAATAGGCGACACGGAATGTTGAATCTCATCTCTT 123
OY 3541 CTTTTCATATTTATTTAGAGCATTTTACAGGGTTATGCTCATGAGCGGATACATTT 3600
Db 122 CTTTTCATATTTATTTAGAGCATTTTACAGGGTTATGCTCATGAGCGGATACATTT 63
OY 3601 TGAATGATTTAGAAAAATTAACAATAGGGGTTCCGCGACATTTCCCGAAAAAGTCC 3660
Db 62 TGAATGATTTAGAAAAATTAACAATAGGGGTTCCGCGACATTTCCCGAAAAAGTCC 3
OY 3661 AC 3662
Db 2 AC 1

```

RESULT 5  
ADM68448/c  
ADM68448 standard; DNA; 3637 BP.

AC ADM68448;  
XX  
DT 03-JUN-2004 (first entry)  
XX  
DE Cycle 3 green fluorescent protein, GFP, DNA construct.  
XX  
KW ds; mismatch endonuclease; endonuclease; gene shuffling technology;  
KW single nucleotide polymorphism; cancer susceptibility;  
KW green fluorescent protein; cycle 3 GFP;  
XX  
OS Aequorea victoria.  
XX  
PN US2003157682-A1.  
XX  
PD 21-AUG-2003.  
XX  
PF 31-JAN-2003; 2003US-00356708.  
XX  
PR 01-FEB-2002; 2002US-0035722P.  
PR 14-MAR-2002; 2002US-00098155.  
PR 01-AUG-2002; 2002US-00211079.  
XX  
PA (PADG/) PADGETT H S.  
PA (VAEW/) VAERHONGS A A.  
PA (VOUD/) VOUDANT F S.  
PA (SMIT/) SMITH M L.  
PA (LIND/) LINDBO J A.  
PA (FITZ/) FITZMAURICE W P.  
XX  
PI Padgett HS, Vaerhongs AA, Vojdani FS, Smith ML, Lindbo JA;  
PI Fitzmaurice WP;  
XX  
DR WPI; 2003-766176/72.  
XX

PT Making a mismatch endonuclease, useful in gene shuffling and in detection  
PT of single nucleotide polymorphisms, comprises transfecting a host with a  
PT recombinant viral vector including a polynucleotide encoding a mismatch  
PT endonuclease.  
XX  
PS Example 9; SEQ ID NO 17; 79pp; English.  
XX  
CC The invention relates to a method of making a mismatch endonuclease  
CC enzyme comprising transfecting a host plant, animal, yeast, fungus or  
CC bacterium with a recombinant viral vector that encodes a polynucleotide  
CC sequence for a mismatch endonuclease, growing the host so that the  
CC polynucleotide is expressed, and extracting the host so that the  
CC enzyme from the host. The method is useful for making mismatch  
CC endonuclease enzymes, for obtaining peptides and polynucleotides with  
CC desired functional properties and for detecting mutations. The mismatch  
CC endonuclease enzymes are useful in gene shuffling technology for  
CC developing new genes, in detecting single nucleotide polymorphisms for  
CC e.g. detecting evidence of cancer susceptibility, or in redistributing  
CC sequence variations between non-identical polynucleotide sequences. The  
CC DNA construct.  
XX  
SQ Sequence 3637 BP; 942 A; 892 C; 870 G; 933 T; 0 U; 0 Other;  
Query Match 72.8%; Score 2664.8; DB 11; Length 3637;  
Best Local Similarity 83.8%; Pred. No. 0;  
Matches 3070; Conservative 0; Mismatches 567; Indels 25; Gaps 4;

OY 1 CTAATTTGTAAGCGTATATATTTGTTAAATTCGGGTTAAATTTGTTAAATTCAGCTC 60  
Db 3637 CTAATTTGTAAGCGTATATATTTGTTAAATTCGGGTTAAATTTGTTAAATTCAGCTC 3578  
OY 61 ATTTTTPAACCAATAGGCGCGAAATCCGCAAAATCCCTTAAATCAAAAGAAATAGACGA 120  
Db 3577 ATTTTTPAACCAATAGGCGCGAAATCCGCAAAATCCCTTAAATCAAAAGAAATAGACGA 3518  
OY 121 GATAGGGTTAGTGTGTTTCCAGTTTGAAACAAGAGTCCATTTAAAGAAAGTGGACTC 180  
Db 3517 GATAGGGTTAGTGTGTTTCCAGTTTGAAACAAGAGTCCATTTAAAGAAAGTGGACTC 3458  
OY 181 CAACGTCAAAAGGCGAAAAACCGTCTATCAGGGCGATGCGCACTGTAACCATCACC 240  
Db 3457 CAACGTCAAAAGGCGAAAAACCGTCTATCAGGGCGATGCGCACTGTAACCATCACC 3398  
OY 241 CTATCAAGTTTGTGGGCTCAGGTGCGTAAAGCACTAAATGGAACCTTAAGGAG 300  
Db 3397 CTATCAAGTTTGTGGGCTCAGGTGCGTAAAGCACTAAATGGAACCTTAAGGAG 3338  
OY 301 CCCCGATTTAGAGCTTGACGGGGAAAGCCGGCAACGTGCGAAGAAAGGAAAGAA 360  
Db 3337 CCCCGATTTAGAGCTTGACGGGGAAAGCCGGCAACGTGCGAAGAAAGGAAAGAA 3278  
OY 361 AGCGAAAGAGCGGGCGCTGAGGGCGCTGCAAGTGTAGCGGTCAACCTGCGCTAAC 420  
Db 3277 AGCGAAAGAGCGGGCGCTGAGGGCGCTGCAAGTGTAGCGGTCAACCTGCGCTAAC 3218  
OY 421 CACACCCCGCGGCTTATGCGCCGCTACAGGGCGGTCCATTCGCATTCAGGCTGCG 480  
Db 3217 CACACCCCGCGGCTTATGCGCCGCTACAGGGCGGTCCATTCGCATTCAGGCTGCG 3158  
OY 481 CAACGTTGGGAAAGGCGATCGGTGCGGGCTTTCGCTATTAACCGACGTGGCAAGG 540  
Db 3157 CAACGTTGGGAAAGGCGATCGGTGCGGGCTTTCGCTATTAACCGACGTGGCAAGG 3098  
OY 541 GGGATGTGCTGAAGGAGATTAATTTGGTTAAGCCAGGGTTTTCCAGTACGACGTTG 600  
Db 3097 GGGATGTGCTGAAGGAGATTAATTTGGTTAAGCCAGGGTTTTCCAGTACGACGTTG 3038  
OY 601 TAAACGACGCGCAAGTGAAGCGCGCTCGTTCACTTCAAGCTTTTGAACCCGTTGAGACG 660  
Db 3037 TAAACGACGCGCAAGTGAAGCGCGCTCGTTCACTTCAAGCTTTTGAACCCGTTGAGACG 2978  
OY 661 GCAGACTCGCGGTAAGTGTGTTTACAGCGGTGATGAGAGATGAAGTGTGACAC 720

```
Db 2977 CCGCGGTGGCGG-----CCGCTCTAAGAACTAGTGATCCCC--CGG 2939
Qy 721 GCTGAGAAACACGACGTAAGATTAACCTAGAAAAGATATATATTTGACGATCGTTAA 780
Db 2938 GCTGAGGAAATTTCTTAATTTGTAATAGTTCATCCATCCCATGTGTAAATCCAGACAGATTA 2879
Qy 781 AGATTAATGATGGTAAAAATTGACGATGGGAATCTGTAAATAGACACTCATATAGGCGCAAT 840
Db 2878 CAAACTCAAGAAAGACCAATGTGTACAGCTTTTGTGTGGAAATCTTTCCAAAGGGCAGATT 2819
Qy 841 TGGGTACCGGGCCCCCTCGAGGTGACAGGTATGATTAAGCTTGATATCGAATTCGATCGC 900
Db 2818 GTGTGACAGGTAATGTGTGTCTGTGTAAAGGACAGGGCCATCCGCCAATTGGAAATTTT 2759
Qy 901 AGCCCGGGGATTCACATGTTCTAGAGCGGCCGCCACCGCGGTGAGCTCCAGCTTTTGT 960
Db 2758 GTTGAATAATGTCTGTCTAGTTGAACGGAATCCATCTTCAATGTGTGGGAAATTTTGAAGT 2699
Qy 961 TCCCTTAGTGAGGGTTAAATAGATCCGATGGGTCAATTTTAAACGACATATCTTTCTAG 1020
Db 2698 TAGCTTAATTCATTTCTTTTGTGTGTCTGTGCGGTGATGTATACATTTGTGTAG--TTATAG 2640
Qy 1021 GATTAACTAGCTGCATCAGATCATATCGTCGGCTCTTTTTCGCGCTCAGTCATCCGC 1080
Db 2639 TTGTACTGAGTTGTGTGTCCGAGAAATGTTTCCATCTTCTTAAATCAATACCTTTTAA 2580
Qy 1081 CAAGTCGGCTATCTGGGCAATCGGGAGAGAAAGCCGTCCTTTTCCGCGAGGTTG 1140
Db 2579 TCGATACATTTAACMAAGGGTATCACCTTCAAACTTGACTTGACGCGCTGTGTAGTTG 2520
Qy 1141 AAGCGCATGAAAAGATTTGCCGAGATGACTGTGTGCAATTTGACCTTGAGCGAANAAC 1200
Db 2519 CCGTCATCTTTGAAAGATATATGTGTCTGTGTATACATCTTCCGCGCATGCACTCTTG 2460
Qy 1201 GCAGTTTACCATGATGATTCGGGAAGGTGTGCGCATGACGCTTTTAAACGGTGAACGT 1260
Db 2459 AAAAAGTATGCCGTTTCAATATGATCCGGAATACGGGAAGGCAATGAACACATPAAG 2400
Qy 1261 TCGTTACAGGCACTGGGATACCAAGTGTGTGGCTTTTCCGACACAGTTCCGATGG 1320
Db 2399 AAAGTAGTGAACAAGTTGTGGCCATGGAACAGGTAGTTTCCAGTAGTGCATAATAATT 2340
Qy 1321 TCAGCCGCAAGCGATTCAGCAACCGGAACATACCGGGACAGCGCGGAACGTGCGTCCG 1380
Db 2339 AGGGTAACCTTTCGTATGTAGCATCACCTTCAACCTCTCCATGACGAAAAATTTTGGC 2280
Qy 1381 GTGTGAGATTAATGACAGCGGTGGCGCTGGGATTTAAGTACAGGAGGACGGGTATC 1440
Db 2279 CCATTAACTACCATCTTAATTCAAACAAGAAATGGGACAACCTCCAGTBAAAAATTTCT 2220
Qy 1441 CTGGCTGATGCCGAGAAATGGAATGGAATACCCCGTAGTTACCCGCGGCGCGCTT 1500
Db 2219 C---CTTTACTCATGCGTAACCAAGCTTTTGTTCCTTTAGTAGAGGATTAATTTGCGCGCTT 2163
Qy 1501 GGCGTAATCATGTCATAGCTGTTTCTGTGTGAAATTTGTTATCCGCTCACAATTTCCCA 1560
Db 2162 GGCGTAATCATGTCATAGCTGTTTCTGTGTGAAATTTGTTATCCGCTCACAATTTCCCA 2103
Qy 1561 CAACATACGAGCCGGAAGCATTAAGTAAAGCTGGGGTGCCTTAATAGTGAAGTAACT 1620
Db 2102 CAACATACGAGCCGGAAGCATTAAGTAAAGCTGGGGTGCCTTAATAGTGAAGTAACT 2043
Qy 1621 CACATTAATTTGCTTGCCTCACTGCGCTTTTCCAGTCCGGAACCTGTGTGCACT 1680
Db 2042 CACATTAATTTGCTTGCCTCACTGCGCTTTTCCAGTCCGGAACCTGTGTGCACT 1983
Qy 1681 GCATTAATGATTCGGCCTAACCGCGGGGAGAGCGGCTTTGGGTATTTGGGCGCTCTTCCG 1740
Db 1982 GCATTAATGATTCGGCCTAACCGCGGGGAGAGCGGCTTTGGGTATTTGGGCGCTCTTCCG 1923
Qy 1741 TTCTCGCTCACTGACTGCTGCGCTCGGTGCTTGGCTGGCGAGGGGTATCAGTCA 1800
Db 1922 TTCTCGCTCACTGACTGCTGCGCTCGGTGCTTGGCTGGCGAGGGGTATCAGTCA 1863
Qy 1801 CTCAAAAGCGGTAATACGGTTATTCACAGAAATCAGGGGATTAACGAGAAAGAACTATG 1860
Db 1862 CTCAAAAGCGGTAATACGGTTATTCACAGAAATCAGGGGATTAACGAGAAAGAACTATG 1803
Qy 1861 AGCAAAAGGCGAGAAAGGCGAGAAACCGTAAGAAAGGCGCGTGTGGCGGTTTTTCA 1920
Db 1802 AGCAAAAGGCGAGAAAGGCGAGAAACCGTAAGAAAGGCGCGTGTGTGGCGGTTTTTCA 1743
Qy 1921 TAGGCTCGCCCCCTGACGAGCATCAAAAATCGACGCTCAAGTCAGAGGTGGGAAA 1980
Db 1742 TAGGCTCGCCCCCTGACGAGCATCAAAAATCGACGCTCAAGTCAGAGGTGGGAAA 1683
Qy 1981 CCCGACAGGACTATTAAGATTCAGGCGTTTTCCCTTGAAGCTCCTGTGTGCGCTTCC 2040
Db 1682 CCCGACAGGACTATTAAGATTCAGGCGTTTTCCCTTGAAGCTCCTGTGTGCGCTTCC 1623
Qy 2041 TGTTCGAGCCCTGCGGCTTACCGGATACGTGTGCGGCTTCTCCCTTCCGGAACGTGGC 2100
Db 1622 TGTTCGAGCCCTGCGGCTTACCGGATACGTGTGCGGCTTCTCCCTTCCGGAACGTGGC 1563
Qy 2101 GCTTCTCATATGCTACAGCTGTAGTATCTCAGTTCCGTGTAGGTGCTTCCCTCAAGCT 2160
Db 1562 GCTTCTCATATGCTACAGCTGTAGTATCTCAGTTCCGTGTAGGTGCTTCCCTCAAGCT 1503
Qy 2161 GGGCTGTGTGACGAACCCCGTTTCAAGCCGACCGCTGCGCTTATTCGGTAACTATCG 2220
Db 1502 GGGCTGTGTGACGAACCCCGTTTCAAGCCGACCGCTGCGCTTATTCGGTAACTATCG 1443
Qy 2221 TCTTGAATCCAACCCGGTAAAGACAGACTTATCGCACCTGGCAGACGCACTGGTAAACAG 2280
Db 1442 TCTTGAATCCAACCCGGTAAAGACAGACTTATCGCACCTGGCAGACGCACTGGTAAACAG 1383
Qy 2281 GATTAGACAGACGAGTATGTAGCGGTGTCTACAGATTTCTTGAAGTGTGGCTTAACTA 2340
Db 1382 GATTAGACAGACGAGTATGTAGCGGTGTCTACAGATTTCTTGAAGTGTGGCTTAACTA 1323
Qy 2341 CGGCTACACTTAAGAAAGACAGTATTTGATTCGTGCTGTGTGAAGCCAGTTACTTCCG 2400
Db 1322 CGGCTACACTTAAGAAAGACAGTATTTGATTCGTGCTGTGTGAAGCCAGTTACTTCCG 1263
Qy 2401 AAAAAGGTTGTATGCTCTTGAATCCGGCAACAAACACCGCTGTAGCGGTGTTTTT 2460
Db 1262 AAAAAGGTTGTATGCTCTTGAATCCGGCAACAAACACCGCTGTAGCGGTGTTTTT 1203
Qy 2461 TGTTTGACAGACAGATTAACGCGCAGAAAAAAGATCTCAAGAAATCTTTGATCTT 2520
Db 1202 TGTTTGACAGACAGATTAACGCGCAGAAAAAAGATCTCAAGAAATCTTTGATCTT 1163
Qy 2521 TTCTACGGGCTGTGACGCTCAAGTGAACGAACCTCAAGTTAAGGAATTTGTGATAG 2580
Db 1142 TTCTACGGGCTGTGACGCTCAAGTGAACGAACCTCAAGTTAAGGAATTTGTGATAG 1083
Qy 2581 ATTATCAAAAAGATCTTCACTCACTAGATCTTTTAAATTAAGAAATTTAATCAAT 2640
Db 1082 ATTATCAAAAAGATCTTCACTCACTAGATCTTTTAAATTAAGAAATTTAATCAAT 1023
Qy 2641 CTTAAAGTATATAGTAAACTTGTGTGACAGTTACCAATGCTTAATCAAGTGAAGCAC 2700
Db 1022 CTTAAAGTATATAGTAAACTTGTGTGACAGTTACCAATGCTTAATCAAGTGAAGCAC 963
Qy 2701 TATCTACGCAATCTGTCTAATTTGTTGTTATCATATAGTTGCTGACTCCCGCTGTAGAT 2760
Db 962 TATCTACGCAATCTGTCTAATTTGTTGTTATCATATAGTTGCTGACTCCCGCTGTAGAT 903
Qy 2761 AACTACGATACGGAGGGCTTACCATCTGAGCCCAAGTGTGCAATGTATCCGAGAACCC 2820
Db 902 AACTACGATACGGAGGGCTTACCATCTGAGCCCAAGTGTGCAATGTATCCGAGAACCC 843
Qy 2821 ACGCTCAACGGCTTCAAGATTTATCAGCAATPAACACGCGGAGGGCGAGCGCAG 2880
Db 842 ACGCTCAACGGCTTCAAGATTTATCAGCAATPAACACGCGGAGGGCGAGCGCAG 783
```



```

QY 2881 AAGTGTCTCGCACTTTATTCGCTCCATCCAGTCTATTATTTGTCGGGAACTAG 2940
DB 782 AAGTGTCTCGCACTTTATTCGCTCCATCCAGTCTATTATTTGTCGGGAACTAG 723
QY 2941 AGTAAGTATTCGCAAGTTAATAGTTTGGCCAAAGTTGTCGCAATGCTACAGGATCGT 3000
DB 722 AGTAAGTATTCGCAAGTTAATAGTTTGGCCAAAGTTGTCGCAATGCTACAGGATCGT 663
QY 3001 GGTGTCAGCTCCGCTGTTGGTATAGCTTCAATCAGCTCCGCTCCAAACATCAAGGCG 3060
DB 662 GGTGTCAGCTCCGCTGTTGGTATAGCTTCAATCAGCTCCGCTCCAAACATCAAGGCG 603
QY 3061 AGTTACATGATCCCGCATGTTGTGCAAAAAGGCTTACCTCTTCGCTCCGATCGT 3120
DB 602 AGTTACATGATCCCGCATGTTGTGCAAAAAGGCTTACCTCTTCGCTCCGATCGT 543
QY 3121 TGTGAGAAATAGTATGCGCCGAGTGTATCACTATGTTATGAGAGACATGCAATATTC 3180
DB 542 TGTGAGAAATAGTATGCGCCGAGTGTATCACTATGTTATGAGAGACATGCAATATTC 483
QY 3181 TCTTACTGTATGCAATCCGTAAGATGCTTTCTGTATGCTGTAGTACTCAACAAGTC 3240
DB 482 TCTTACTGTATGCAATCCGTAAGATGCTTTCTGTATGCTGTAGTACTCAACAAGTC 423
QY 3241 ATTCTGAGAAATAGTATGCGCCGAGTGTATCACTATGTTATGAGAGACATGCAATATTC 3300
DB 422 ATTCTGAGAAATAGTATGCGCCGAGTGTATCACTATGTTATGAGAGACATGCAATATTC 363
QY 3301 TACCCGCGCAATAGCAAACTTTAAAGTCTATCATCTGAAAGAGTTCTTCTCGGCGC 3360
DB 362 TACCCGCGCAATAGCAAACTTTAAAGTCTATCATCTGAAAGAGTTCTTCTCGGCGC 303
QY 3361 AAAACTCTCAAGATCTTACCGCTGTGAGATCAAGTTGATGATCAACCACTCGTGACC 3420
DB 302 AAAACTCTCAAGATCTTACCGCTGTGAGATCAAGTTGATGATCAACCACTCGTGACC 243
QY 3421 CAATGATCTTCAAGATCTTATCTTCAACGAGCTTCTGCTGAGTACCACTCGTGACC 243
DB 242 CAATGATCTTCAAGATCTTATCTTCAACGAGCTTCTGCTGAGTACCACTCGTGACC 183
QY 3481 GCAAAATGCGCAAAAAGGAAATAGGCGGACACGAAATGTTGAATACATCTCTT 3540
DB 182 GCAAAATGCGCAAAAAGGAAATAGGCGGACACGAAATGTTGAATACATCTCTT 123
QY 3541 CTTTTCATATTAATGAAATTAATCAAGATTAATGAGGTTATGCTATGAGGATACATAT 3600
DB 122 CTTTTCATATTAATGAAATTAATCAAGATTAATGAGGTTATGCTATGAGGATACATAT 63
QY 3601 TGAATGATTTAGAAAATTAACAATAAGGCTTCCGCGCAATTTCCCGAAAAGTGC 3660
DB 62 TGAATGATTTAGAAAATTAACAATAAGGCTTCCGCGCAATTTCCCGAAAAGTGC 3
QY 3661 AC 3662
DB 2 AC 1

```

```

RESULT 6
ADP26601/c
ID ADP26601 standard; DNA; 3637 BP.
XX
AC ADP26601;
XX
DT 26-AUG-2004 (first entry)
XX
DE Green fluorescent protein (GFP) plasmid DNA #2.
XX
KM Sequence variation; heteroduplex; transcription; DNA integration;
XX
KW ribozyme expression; gene; ds; green fluorescent protein; GFP.
OS Aequorea victoria.
XX
Synthetic.

```

```

XX
PN US2004110130-A1.
XX
XX 10-JUN-2004.
XX
PF 25-OCT-2002; 2002US-00280913.
XX
PR 02-FEB-2001; 2001US-0266386P.
PR 14-FEB-2001; 2001US-0268785P.
PR 01-FEB-2002; 2002US-00066390.
PR 08-AUG-2002; 2002US-0402342P.
XX
PA (LARG-) LARGE SCALE BIOLOGY CORP.
XX
PI Padgett HS, Lindbo JA, Fitzmaurice WP,
XX
DR WPI, 2004-440326/41.
XX
PT Redistributing sequence variations between non-identical polynucleotide
XX
PT sequences, useful for generating improved polynucleotide having a desired
XX
PT characteristic, comprises making a heteroduplex and introducing a nick.
XX
PS Example 5; SEQ ID NO 17; 75pp; English.
XX
CC The invention relates to an in vitro method of redistributing sequence
XX
CC variations between non-identical polynucleotide sequences, comprising
XX
CC making a heteroduplex polynucleotide from two non-identical
XX
CC polynucleotides, introducing a nick in the second strand at or near a
XX
CC base pair mismatch site, removing the mismatched base(s) from the
XX
CC mismatch site where the nick occurred and using the first strand as a
XX
CC template to replace the removed base(s) with bases that complement the
XX
CC base(s) in the first strand. The invention also relates to an in vitro
XX
CC method of making a population of sequence variants from a heteroduplex
XX
CC polynucleotide sequence, obtaining a polynucleotide sequence encoding a
XX
CC desired functional property and identifying a polynucleotide sequence
XX
CC encoding a protein with an improved functional property. The method is
XX
CC of improved polynucleotide sequences possessing at least one desired
XX
CC phenotypic characteristic (e.g., promotes transcription of linked
XX
CC polynucleotides), where said polynucleotides are useful for expression
XX
CC from a plant, animal, fungal, yeast, or bacterial expression vector, for
XX
CC integration to form a transgenic plant, animal or microorganism, and for
XX
CC expression of a ribozyme. This sequence represents DNA used in the scope
XX
XX
SQ Sequence 3637 BP; 942 A; 892 C; 870 G; 933 T; 0 U; 0 Other;

```

Query Match 72.8%; Score 2664.8; DB 12; Length 3637;  
 Best Local Similarity 83.8%; Pred. No. 0;  
 Matches 3070; Conservative 0; Mismatches 567; Indels 25; Gaps 4;

```

QY 1 CTAATTTGTAAGCGTTAATATTTGTTAAATTCGCTTAATTTGTTAATAGCTC 60
DB 3637 CTAATTTGTAAGCGTTAATATTTGTTAAATTCGCTTAATTTGTTAATAGCTC 3578
QY 61 ATTTTAAACCAATAGGCGGAATCGGCAAAATCCCTTAATTAATCAAAAGATGACCGA 120
DB 3577 ATTTTAAACCAATAGGCGGAATCGGCAAAATCCCTTAATTAATCAAAAGATGACCGA 3518
QY 121 GATAGGTTAGTGTGTTCCAGTTTGGAAACAAGATCCATTAATTAAGAACTGGAATC 180
DB 3517 GATAGGTTAGTGTGTTCCAGTTTGGAAACAAGATCCATTAATTAAGAACTGGAATC 3458
QY 181 CAACGTCAAAAGGCGAAAACGCTATCAAGGCGGATGCGCACTAGTGAACCATACACC 240
DB 3457 CAACGTCAAAAGGCGAAAACGCTATCAAGGCGGATGCGCACTAGTGAACCATACACC 3398
QY 241 CTAAATAGCTTTTGGGCTGAGGTGCGTAAAGCACTAAATCGAAACCTTAAAGGAG 300
DB 3397 CTAAATAGCTTTTGGGCTGAGGTGCGTAAAGCACTAAATCGAAACCTTAAAGGAG 3338
QY 301 CCCCAGTTTGAAGCTTGAACGGGGAAGCCGGCAACGTGCGCAAGAAAGAGGAAGAA 360

```



Dh 3337 CCCCCGATTTAGAGCTTGACGCGGGAAAGCCGGCGAAGCTGGCGAAGAAAGGAGAA 3278  
Qy |||||  
Dh 361 AGCGAAAGAGCGCGCTTAGGGCCCTGGCAAGTGTAGCGGTCAAGCTGCGGTAACAC 420  
Qy |||||  
Dh 3277 AGCGAAAGAGCGCGCTTAGGGCCCTGGCAAGTGTAGCGGTCAAGCTGCGGTAACAC 3218  
Qy |||||  
Dh 421 CACACCCCGCGCTTAAATGCGCGCTACAGGGCGCGTCCCATTCGCCATTCAGGCTGG 480  
Qy |||||  
Dh 3217 CACACCCCGCGCTTAAATGCGCGCTACAGGGCGCGTCCCATTCGCCATTCAGGCTGG 3158  
Qy |||||  
Dh 481 CAACGTGTGGAAAGGCGCATCGGTGCGGGCTCTTCGCTATTACGCCAGCTGGCGAAG 540  
Qy |||||  
Dh 3157 CAACGTGTGGAAAGGCGCATCGGTGCGGGCTCTTCGCTATTACGCCAGCTGGCGAAG 3098  
Qy |||||  
Dh 541 GGGATGTCTGCAAGCGCATTAAGTTGGGTAAAGCCAGAGGTTTTCCAGTCAAGCTGG 600  
Qy |||||  
Dh 3097 GGGATGTCTGCAAGCGCATTAAGTTGGGTAAAGCCAGAGGTTTTCCAGTCAAGCTGG 3038  
Qy |||||  
Dh 601 TAAACGACGCGCACTGAGCGCGCTCGTTCATTCACGTTTTTGAACCGGTGAGAGAG 660  
Qy |||||  
Dh 3037 TAAACGACGCGCACTGAGCGCGCTCGTTCATTCACGTTTTTGAACCGGTGAGAGAG 2978  
Qy |||||  
Dh 661 GCAGACTGCGGTGCAAAATGTGTATTACAGCGTGAATGAGCAGATGAAGATGCTGAGAC 720  
Qy |||||  
Dh 2977 CCGCGGTGCGG-----CGCTCTAGAACTAGTGGATCCC--CGG 2939  
Qy |||||  
Dh 721 GCTGAGAACACGCACTAGATTAACCTAGAAATATATCATATTGTGACGTACGTAA 780  
Qy |||||  
Dh 2938 GCTGAGAAATCTTATTTGTATAGTTTATTCATTCATTCATTCATTCATTCATTCATTC 2879  
Qy |||||  
Dh 781 AGATATCATGCTGTAATTTGACGATGGAGCTGTAAATGACACTCATATAGGCGAGT 840  
Qy |||||  
Dh 2878 CAATCTCAAGAAAGACATGTGTGTCACGCTTTTCGTGGAGCTTTGAAAGGAGAGATT 2819  
Qy |||||  
Dh 841 TGGGTACCGGCGCCCTCGAGGTGACGATGATGATAGCTTGATATGCAATTCCTG 900  
Dh 2818 GTGTGACAGGTATGTTGTGTCTGTGTAAGAACAGAGGCAATCGCAATTGGAGATTTT 2759  
Qy |||||  
Dh 901 AGCCGCGGAGATCCATAGTTCTAGAGCGCGCCACCGCGGTGAGCTCCAGCTTTTGT 960  
Qy |||||  
Dh 2758 GTTGTAAATGCTGCTAGTTGAAAGCATTCATTCATTCATTCATTCATTCATTCATTC 2699  
Qy |||||  
Dh 961 TCCCTTAGTGAAGGTATTAATAGATCCATGCGTCAATTTTACGACATATCTTTAG 1020  
Qy |||||  
Dh 2698 TAGCTTATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 2640  
Qy |||||  
Dh 1021 GGTATATCTAGCTGATCAGATCATATCGTGGGCTTTTTCGCGCTCAGTCAATCCG 1080  
Dh 2639 TTGTACTGAGTTGTGTGCGAGATGTTTCCATCTTCTTAAATCAATACCTTTTAA 2580  
Qy |||||  
Dh 1081 CAAGCTGCGCTATCTGGGATCGGGAGAAAGAACCCGTGCTTTTCCGCGAGTTG 1140  
Qy |||||  
Dh 2579 TCGATACATTAACAAGGATACACCTTCAAACTTGACTTGACGCGGTCTGTGATG 2520  
Qy |||||  
Dh 1141 AAGCGCATGAGAAAGTTTCCGAGATGACGTGCTGATTCATTCATTCATTCATTCATTC 1200  
Dh 2519 CCGTCATCTTGAAGATATAGTGTCTGTATATAACCTTCGAGATGCACTTTG 2460  
Qy |||||  
Dh 1201 GCAGCTTACATGATGATTCGGAGAGTGTGGCCATGACGCTTTTAAAGGTGAATGT 1260  
Dh 2459 AAAAGTATGATGCTTATATGATCCGGAATACGGAAAGCATTTGAACACATPAAG 2400  
Qy |||||  
Dh 1261 TCGTTACGCGCACTGGGATACAGTGTGTGCGGCTTTTCCGACACAGTTCCGATGG 1320  
Dh 2399 AAAGTATGACAAAGTGTGGCATGGAACAGGTATTTTCCAGTGTGCTCAATTAATTA 2340  
Qy |||||  
Dh 1321 TCAGCCGAGAGCATGACAAACCGGAACATACCGGGGACGCGCGGAACTGCGTGGC 1380  
Dh 2339 AGGTATGCTTTCGTATGATGATCACCCTTCCATGACGAAATTTGTGC 2280  
Qy |||||  
Dh 1381 GTGTGACATTAATGACAGCGGTGCGGCTGGGATATTACGTACGAGAGCGGGTATC 1440  
Dh 2279 CCATTAACATGACATCTTAATTCACAAAGAAATGGGAGCACTCGAGTGAAGTTCTTCT 2220

Qy 1441 CTGCTGATGATCCCGCAAAATGACATGATATCCCGGTGAGTTACCGGCGGCGCGCTT 1500  
Dh C---CTTATCTCATCGGTATCCCACTTTTGTCTCTTATGAGGGTTAATTTGGCGCTT 2163  
Qy 1501 GCGGTAATCATGATCATATGATGTTTCTGTGTGAATTTGTTATCGCTCACAATTCACA 1560  
Dh 2162 GCGGTATCATGATCATATGATGTTTCTGTGTGAATTTGTTATCGCTCACAATTCACA 2103  
Qy 1561 CAACATACGACCGGAGCATTAAGTGAAGCTGGGCTGCTTAATGATGATTAATCT 1620  
Dh 2102 CAACATACGACCGGAGCATTAAGTGAAGCTGGGCTGCTTAATGATGATTAATCT 2043  
Qy 1621 CACATTAATTCGTTGGCTCTACCTGCTTTTCCAGTGGGAAACCTGTGTGCACT 1680  
Dh 2042 CACATTAATTCGTTGGCTCTACCTGCTTTTCCAGTGGGAAACCTGTGTGCACT 1983  
Qy 1681 GCATTAATGATTCGCGCAACGCGGGGAGAGCGGTTTGGGTATTGGGCGCTCTTCCGC 1740  
Dh 1982 GCATTAATGATTCGCGCAACGCGGGGAGAGCGGTTTGGGTATTGGGCGCTCTTCCGC 1923  
Qy 1741 TTCTCGCTCATGATCTGCTGCGCTGCTGCTTCCGCTGCGGAGCGGATATCAGCTCA 1800  
Dh 1922 TTCTCGCTCATGATCTGCTGCGCTGCTGCTTCCGCTGCGGAGCGGATATCAGCTCA 1863  
Qy 1801 CTCAAAAGCGGTATATCGGTATATCCACAGATTCAGGGGATTAAGCGAAGAAACATGTG 1860  
Dh 1862 CTCAAAAGCGGTATATCGGTATATCCACAGATTCAGGGGATTAAGCGAAGAAACATGTG 1803  
Qy 1861 AGCAAAAGGCGCAGCAAAAGGCGCAGAAACCGTAAAGGCGCGTGTGCTGGCGTTTTC 1920  
Dh 1802 AGCAAAAGGCGCAGCAAAAGGCGCAGAAACCGTAAAGGCGCGTGTGCTGGCGTTTTC 1743  
Qy 1921 TAGCTCGCGCCCTGACGAGATCACAAAATTCAGCTCAAGTCAAGAGTGGCGAAA 1980  
Dh 1742 TAGCTCGCGCCCTGACGAGATCACAAAATTCAGCTCAAGTCAAGAGTGGCGAAA 1683  
Qy 1981 CCGCAGAGATTAATTAAGATTCAGAGGCTTCCCTCGAGAGCTCCCTCGGCTCTCC 2040  
Dh 1682 CCGCAGAGATTAATTAAGATTCAGAGGCTTCCCTCGAGAGCTCCCTCGGCTCTCC 1623  
Qy 2041 TGTTCGACCTGCGCTTACCGGATACGTGTGCGGCTTTCCTTCGGAACGATGAGC 2100  
Dh 1622 TGTTCGACCTGCGCTTACCGGATACGTGTGCGGCTTTCCTTCGGAACGATGAGC 1563  
Qy 2101 GCTTCTCATATGATCAAGCTGTATGATCTTCAAGTTGCGGTGATGCTTCCCAAGCT 2160  
Dh 1562 GCTTCTCATATGATCAAGCTGTATGATCTTCAAGTTGCGGTGATGCTTCCCAAGCT 1503  
Qy 2161 GGGCTGTGTGACGAACCCCGCTTACAGCCGACCGCTGCGCTTATCCGTTAACTATCG 2220  
Dh 1502 GGGCTGTGTGACGAACCCCGCTTACAGCCGACCGCTGCGCTTATCCGTTAACTATCG 1443  
Qy 2221 TCTTGAATGCAACCGGATTAAGACAGCTTATGCGCATGCGCAGCAGCACTGGTAA 2280  
Dh 1442 TCTTGAATGCAACCGGATTAAGACAGCTTATGCGCATGCGCAGCAGCACTGGTAA 2183  
Qy 2281 GATTAGAGAGAGGATATGAGCGGTGCTACAGATTTCTTGAAGTGTGGCTTAACTA 2340  
Dh 1382 GATTAGAGAGAGGATATGAGCGGTGCTACAGATTTCTTGAAGTGTGGCTTAACTA 1323  
Qy 2341 CCGCTACATTAAGAGACATATTTGTTATGCTGCTGTGTGAAGCAAGTTACCTTGG 2400  
Dh 1322 CCGCTACATTAAGAGACATATTTGTTATGCTGCTGTGTGAAGCAAGTTACCTTGG 1263  
Qy 2401 AAAAGATGTTGTTGCTTGTATCGGCAAAACAACACCGCTGTGAGCGGTTTCTT 2460  
Dh 1262 AAAAGATGTTGTTGCTTGTATCGGCAAAACAACACCGCTGTGAGCGGTTTCTT 1203  
Qy 2461 TGTTCGACAGCAGATTAAGCGCAGAAAAAAGATCTCAAGAAATCTTTGATCTT 2520  
Dh 1202 TGTTCGACAGCAGATTAAGCGCAGAAAAAAGATCTCAAGAAATCTTTGATCTT 1143

Qy	2521	TTCTACGGGGCTGAGCGCTCAGCGTGAAGCAAGAAAACTCAGCTTAAGGATTTTGGTCATGAG	2580
Db	1142	TTCTACGGGGGTCTGACGCTCAGCGTGAAGCAAGAAAACTCAGCTTAAGGATTTTGGTCATGAG	1083
Qy	2581	ATTAATCAAAAAGGATTTTCACTTAATCTTTTAAATTTAAATTAATTAATCAAT	2640
Db	1082	ATTATCAAAAAGGATTTTCACTTAATCTTTTAAATTTAAATTAATTAATCAAT	1023
Qy	2641	CTTAAGTATATATGATGAACCTTGGCTGCAGAGTTACCAATGCTTAAATCAATGAGGACC	2700
Db	1022	CTAAAGTATATATGATGAACCTTGGCTGCAGAGTTACCAATGCTTAAATCAATGAGGACC	963
Qy	2701	TATCTCAGCGATCTGTCTAATTCCTTTCCTTCAATCCATAGTTGCTGACTCCCGGTGATAGAT	2760
Db	962	TATCTCAGCGATCTGTCTAATTCCTTTCCTTCAATCCATAGTTGCTGACTCCCGGTGATAGAT	903
Qy	2761	AACATTCGATACGGGAGGGCTTACCATCTGGGCCAGTCTGCAATGATACCGGAGACC	2820
Db	902	AACATTCGATACGGGAGGGCTTACCATCTGGGCCAGTCTGCAATGATACCGGAGACC	843
Qy	2821	ACGCTCAACGGGCTCCAAATTTATCAAGCATTAACCAAGCAGCGAGCGAGGCGAGCGAG	2880
Db	842	ACGCTCAACGGGCTCCAAATTTATCAAGCATTAACCAAGCAGCGAGCGAGGCGAGCGAG	783
Qy	2881	AAGTGGCTCTGCAACTTATCCGCTCCATCCAGTCTATTAATTTGTCGGGAAAGCTAG	2940
Db	782	AAGTGGCTCTGCAACTTATCCGCTCCATCCAGTCTATTAATTTGTCGGGAAAGCTAG	723
Qy	2941	AGTAAGTATAGTTCGCGATTAAATAGTTGGCAAGCTTGTGCAATGCTTACAGGCATCGT	3000
Db	722	AGTAAGTATAGTTCGCGATTAAATAGTTGGCAAGCTTGTGCAATGCTTACAGGCATCGT	663
Qy	3001	GGGTTCACGCTGTGCTGGTTGGTATAGGCTTCAATYAGCTCCGGTCCCAAGATCAAGCG	3060
Db	662	GGGTTCACGCTGTGCTGGTTGGTATAGGCTTCAATYAGCTCCGGTCCCAAGATCAAGCG	603
Qy	3061	AGTTACATGATCCCCCATGTTGTGTGCAAAAAACGCGTTAGCTCTTCGCTCCGATCGT	3120
Db	602	AGTTACATGATCCCCCATGTTGTGTGCAAAAAACGCGTTAGCTCTTCGCTCCGATCGT	543
Qy	3121	TGTCAAGATAGTTTGGCCGAGTGTATCACTCATAGGTTAATGGCAGCATCAATATTC	3180
Db	542	TGTCAAGATAGTTTGGCCGAGTGTATCACTCATAGGTTAATGGCAGCATCAATATTC	483
Qy	3181	TCCTACTGTCATGCGCATCCGTAAGATCCTTTTCTGTGACTGTGAGTACTCAACCAAGTC	3240
Db	482	TCCTACTGTCATGCGCATCCGTAAGATCCTTTTCTGTGACTGTGAGTACTCAACCAAGTC	423
Qy	3241	ATTCTGGAATATGTTATGCGGAGCAGCGAGTTCCTTGGCCGGGGTCAATACGGGATTA	3300
Db	422	ATTCTGGAATATGTTATGCGGAGCAGCGAGTTCCTTGGCCGGGGTCAATACGGGATTA	363
Qy	3301	TACCGCGCCATACGAGCAACTTTAAAGTGTCTCATCATTTGAAAACTTTCGCGGGCG	3360
Db	362	TACCGCGCCATACGAGCAACTTTAAAGTGTCTCATCATTTGAAAACTTTCGCGGGCG	303
Qy	3361	AAAACTCTCAAGATTTTACGCGCTTGTGAGATCCAGTTGATGTAACCACTGTCGACC	3420
Db	302	AAAACTCTCAAGATTTTACGCGCTTGTGAGATCCAGTTGATGTAACCACTGTCGACC	243
Qy	3421	CAACTGATCTTCAGCATCTTTTACTTTTACCAAGCGTTCTTGGGTGACAAAAACAGGAG	3480
Db	242	CAACTGATCTTCAGCATCTTTTACTTTTACCAAGCGTTCTTGGGTGACAAAAACAGGAG	183
Qy	3481	GCAAAATGCCCAAAAAAGGGAATTAAGGCGCAGCAGGAAATGTTGAAATCTCATACTCTT	3540
Db	182	GCAAAATGCCCAAAAAAGGGAATTAAGGCGCAGCAGGAAATGTTGAAATCTCATACTCTT	123
Qy	3541	CCTTTTCATATATATGAAGCATTTATCAAGGTTATTTGCTCATAGCGGATACATATT	3600
Db	122	CCTTTTCATATATATGAAGCATTTATCAAGGTTATTTGCTCATAGCGGATACATATT	63
Qy	3601	TGATGATTTTAAAAAAATTAACAAATAGGGGTTTCGGCCACATTTCCCGAAAAATGCC	3660

Dd	62	TGATGTATTAGAAAAATAAACAAATAGGGGTTCCGGCACAATTTCCTCGAAAAGTGC
Oy	3661	AC 3662
Dd	2	AC 1
RESULT 7		
ABX14570/c		
ID	ABX14570	standard; DNA; 3637 BP.
XX		
AC	ABX14570;	
XX		
DT	14-MAR-2003	(first entry)
DE	Plasmid pBSWTGFP DNA.	
XX		
KV	Circular; cyclic; heteroduplex; T4 DNA polymerase; T4 DNA ligase; CEL I;	
XX	population diversity; exonuclease; complementarity; de.	
OS	Synthetic.	
XX		
FN	M0200279468-A2.	
PD	10-OCT-2002.	
PF	01-FEB-2002; 2002MO-USO03055.	
PR	02-FEB-2001; 2001US-0266386P.	
XX	14-FEB-2001; 2001US-0268785P.	
PA	(LARG-) LARGE SCALE BIOLOGY CORP.	
XX		
Pdgett HS,	Fitzmaurice WP, Lindo JA,	
DR	WPI, 2003-092898/08.	
PT		
PT	Increasing complementarity by making sequence variants from heteroduplex	
XX	polynucleotides, useful for producing improved gene products from	
XX	randomly mutated genes.	
Example 5;	Page 98-99; 105pp; English.	
XX		
XX	The invention relates to an in vitro method of making sequence variants	
CC	from at least one heteroduplex polynucleotide that has at least two-non	
CC	complementary nucleotide base pairs. The method comprises combining a	
CC	heteroduplex polynucleotide with CEI. I, T4 DNA polymerase and T4 DNA	
CC	ligase, or an agent or agents with exonuclease activity and allowing	
CC	sufficient time for the percentage of complementarity to increase, where	
CC	one or more variants are made. The method can also be used to increase	
CC	diversity in a population of sequences. The method is useful for	
CC	producing improved gene products from randomly mutated genes or from	
CC	expression from a suitable plant, animal, fungal, yeast or bacterial	
CC	recombinants for generating novel polynucleotides from parental templates	
CC	that are more diverse and with a lower percentage of sequence identity.	
CC	This sequence represents plasmid DNA, used in heteroduplex substrate	
CC	preparation	
XX		
Sequence 3637 BP;	959 A; 879 C; 868 G; 931 T; 0 U; 0 Other;	
Query Match		
Best Local Similarity	72.5%; Score 2656.4; DB 8; Length 3637;	
Matches 3079;	Conservative 0; Mismatches 556; Indels 29; Gaps 6;	
OY	1 CTAAATGTGAAGCGTTAATATTTTGTTAAATTCGGTTAAATTTTGTTAATCAAGTGC	60
DB	3637 CTAAATGTGAAGCGTTAATATTTTGTTAAATTCGGTTAAATTTTGTTAATCAAGTGC	3578
OY	61 ATTTTAAACCAATAGGCCGAATTCGCAAAATCCCTATTAATCAAAAGAATAGACCGA	120
DB	3577 ATTTTAAACCAATAGGCCGAATTCGCAAAATCCCTATTAATCAAAAGAATAGACCGA	3518

OY 121 GATAGGTTGAGTGTGTTCCAGTTTGGAAACAAGATCCACTTTTAAAGACGTGACATC 180  
DB 3517 GATAGGTTGAGTGTGTTCCAGTTTGGAAACAAGATCCACTTTTAAAGACGTGACATC 3458  
OY 181 CAACGTCAAGGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACATCAAC 240  
DB 3457 CAACGTCAAGGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACATCAAC 3398  
OY 241 CTAAATCAAGTTTTTTGGGGTGCAGGTGCCGTAAAGCACTAAATCCGAACCTTAAAGGAG 300  
DB 3397 CTAAATCAAGTTTTTTGGGGTGCAGGTGCCGTAAAGCACTAAATCCGAACCTTAAAGGAG 3338  
OY 301 CCCCCGATTTTAAAGCTTGAACGGGGAACCGGCGGAACGTGGCGAAGAAAGGAAGGA 360  
DB 3337 CCCCCGATTTTAAAGCTTGAACGGGGAACCGGCGGAACGTGGCGAAGAAAGGAAGGA 3278  
OY 361 AGCGAAAGAGCGGGCGCTAGGGCGCTGGCAAGTGAAGCGATCAAGCTGCGGCTAACAC 420  
DB 3277 AGCGAAAGAGCGGGCGCTAGGGCGCTGGCAAGTGAAGCGATCAAGCTGCGGCTAACAC 3218  
OY 421 CACACCGCGCGGCTTAAATGCGCGCTACAGGGCGCGTCCATTGCGCAATTCAGGCTGCG 480  
DB 3217 CACACCGCGCGGCTTAAATGCGCGCTACAGGGCGCGTCCATTGCGCAATTCAGGCTGCG 3158  
OY 481 CAACGTGTGGGAAGGGCGCATGGGTGCGGCGCTCTTCCCTATTACCGCAGCTGGCGAAAG 540  
DB 3157 CAACGTGTGGGAAGGGCGCATGGGTGCGGCGCTCTTCCCTATTACCGCAGCTGGCGAAAG 3098  
OY 541 GGGATGTGCTGCAAGCGCATTAAGTTGGGTAAACGACAGGGTTTTCCAGTCAAGAGCTTG 600  
DB 3097 GGGATGTGCTGCAAGCGCATTAAGTTGGGTAAACGACAGGGTTTTCCAGTCAAGAGCTTG 3038  
OY 601 TAAAAAGACGCGCACTGAGCGCGCTCGTTCAATTCAGCTTTTGAACCCGTGGAGAGCG 660  
DB 3037 TAAAAAGACGCGCACTGAGCGCGCTCGTTCAATTCAGCTTTTGAACCCGTGGAGAGCG 2978  
OY 661 GCAGACTGGCGGTGCAAAATGTGTTTTTACAGCTGATGAGCAGATGAAGATGCTGACAC 720  
DB 2977 GCAGACTGGCGGTGCAAAATGTGTTTTTACAGCTGATGAGCAGATGAAGATGCTGACAC 2918  
OY 721 GCTGAGAAACAGCAGCTAGATTAAACCTAGAAAGATATCATATTGTGACGTAGCTTAA 780  
DB 2917 GCTGAGAAACAGCAGCTAGATTAAACCTAGAAAGATATCATATTGTGACGTAGCTTAA 2858  
OY 781 AGATATATC -ATGCGTAAATTTGACGATGGATCTGTAAATGCACTGACTATAGAGCGA 838  
DB 2857 AGATATATC -ATGCGTAAATTTGACGATGGATCTGTAAATGCACTGACTATAGAGCGA 2798  
OY 839 ATTGGGTACCGGGCCCCCTGAGAGTGAACGATTCGATTAAGCTTGAATTCGAATTCCT 898  
DB 2797 ATTGGGTACCGGGCCCCCTGAGAGTGAACGATTCGATTAAGCTTGAATTCGAATTCCT 2738  
OY 899 GCAGCCCGGGGGATCCACTAGTTCTAGAGCGCGCGCAACCGCGGTGAGCTCCAGCTTT 958  
DB 2737 GCAGCCCGGGGGATCCACTAGTTCTAGAGCGCGCGCAACCGCGGTGAGCTCCAGCTTT 2678  
OY 959 GTTCCCTTAAAGAGGTTAAATTAAGATCCATGCGCTCAATTTTACGACACTATCTTCT 1018  
DB 2677 GTTCCCTTAAAGAGGTTAAATTAAGATCCATGCGCTCAATTTTACGACACTATCTTCT 2619  
OY 1019 AGGGTTAATCTAGCTGACATCAGATCATATCGTCCGGTCTTTTTCGGGCTCAGTATCG 1078  
DB 2618 AGGGTTAATCTAGCTGACATCAGATCATATCGTCCGGTCTTTTTCGGGCTCAGTATCG 2564  
OY 1079 CCAAGCTGGCGCTATCTGGGATGCGGGAAGGAAGAACCCGTCCTTTCCCGGAGGT 1138  
DB 2563 CCAAGCTGGCGCTATCTGGGATGCGGGAAGGAAGAACCCGTCCTTTCCCGGAGGT 2515  
OY 1139 TGAAGCGGATGAAAGATTTGCGAGAGATGACTGCTGCTGACTTGAAGCTTGAAGCGAA 1198  
DB 2514 TGAAGCGGATGAAAGATTTGCGAGAGATGACTGCTGCTGACTTGAAGCTTGAAGCGAA 2455

OY 1199 AGCAGCTTTACATGATGATTCGGGAAGGTGTGCGCATGACCGCTTAAACGTGAATC 1258  
DB 2454 AGCAGCTTTACATGATGATTCGGGAAGGTGTGCGCATGACCGCTTAAACGTGAATC 2395  
OY 1259 GTTCCGTAGGCGCACTGGGATTAACAGTTGCTGCGGCTTTTTCGGACACAGTTCCGAT 1318  
DB 2394 GTTCCGTAGGCGCACTGGGATTAACAGTTGCTGCGGCTTTTTCGGACACAGTTCCGAT 2335  
OY 1319 GGTCAAGCCGGAAGCGCATCAGCAACCCGAAATACCGGCGACAGCCGGAACCTGCGTGC 1378  
DB 2334 GGTCAAGCCGGAAGCGCATCAGCAACCCGAAATACCGGCGACAGCCGGAACCTGCGTGC 2280  
OY 1379 CGGTGTGAGATTATGACAGCGGTGCGGCGCTGGATATTACGTGACGAGAGCGGTA 1438  
DB 2279 CGGTGTGAGATTATGACAGCGGTGCGGCGCTGGATATTACGTGACGAGAGCGGTA 2220  
OY 1439 TCTGGCTGAGATGCGCGAAGATGAGCATGATACCCCGTGAAGTTACCGGCGGCGCGC 1498  
DB 2219 TCTGGCTGAGATGCGCGAAGATGAGCATGATACCCCGTGAAGTTACCGGCGGCGCGC 2165  
OY 1499 TTGGGTAATCATGGTCAATGACTGTTTCTGTGTGAATTTGTATCCGCTCAAAATTC 1558  
DB 2164 TTGGGTAATCATGGTCAATGACTGTTTCTGTGTGAATTTGTATCCGCTCAAAATTC 2105  
OY 1559 CACAAATACAGAGCCGGAAGCATTAAGTGTAAAGCTTGGGCTGCTAATGAGTGAAT 1618  
DB 2104 CACAAATACAGAGCCGGAAGCATTAAGTGTAAAGCTTGGGCTGCTAATGAGTGAAT 2045  
OY 1619 CTCAATTAATTTGGGTTGCGCTCACTGCGCGCTTTTCCAGTGGGGAACCTGTCGAGCAG 1678  
DB 2044 CTCAATTAATTTGGGTTGCGCTCACTGCGCGCTTTTCCAGTGGGGAACCTGTCGAGCAG 1985  
OY 1679 CTGCAATTAATGAGTACGCGCAACCGCGGGAAGAGCGGTTTGCATATTTGGCGCTTCC 1738  
DB 1984 CTGCAATTAATGAGTACGCGCAACCGCGGGAAGAGCGGTTTGCATATTTGGCGCTTCC 1925  
OY 1739 GCTTCTGCTCACTGACTCGCTGCGCTCGGTGTTGGCTGCGGCGGAGCGGTATCACT 1798  
DB 1924 GCTTCTGCTCACTGACTCGCTGCGCTCGGTGTTGGCTGCGGCGGAGCGGTATCACT 1865  
OY 1799 CACTCAAAAGCGGTATACGCTTATCCACAGAAATCAGGGGATTAACGCGAAGAAACATG 1858  
DB 1864 CACTCAAAAGCGGTATACGCTTATCCACAGAAATCAGGGGATTAACGCGAAGAAACATG 1805  
OY 1859 TGAGCAAAAGCGCAGCAAAAGCGCAAGAACCGTAAAGAGCGGCTGCTGCTTTTC 1918  
DB 1804 TGAGCAAAAGCGCAGCAAAAGCGCAAGAACCGTAAAGAGCGGCTGCTGCTTTTC 1745  
OY 1919 CATAGGCTCGGCCCCCTGACAGCATCAAAAATGACGCTCAAGTCAAGTGGCGA 1978  
DB 1744 CATAGGCTCGGCCCCCTGACAGCATCAAAAATGACGCTCAAGTCAAGTGGCGA 1685  
OY 1979 AACCGACAGGACTATAAGATACAGGCGTTTCCCGTGAAGTCCCTGTCGCTCT 2038  
DB 1684 AACCGACAGGACTATAAGATACAGGCGTTTCCCGTGAAGTCCCTGTCGCTCT 1625  
OY 2039 CCTGTTCGACCCCTGCGCTTACCGGATACCTGTGCGCTTTCTCCCTCGGGAAGCTG 2098  
DB 1624 CCTGTTCGACCCCTGCGCTTACCGGATACCTGTGCGCTTTCTCCCTCGGGAAGCTG 1565  
OY 2099 GCGCTTTCTCATGACTCAAGCTGTAGATATCTCAGTTGCGGTGAGTTCGCTCAAG 2158  
DB 1564 GCGCTTTCTCATGACTCAAGCTGTAGATATCTCAGTTGCGGTGAGTTCGCTCAAG 1505  
OY 2159 CTGGGCTGTGACAGAAACCCCGTTACGCGGCAACGCTGCGCTTATCCGGAATCAT 2218  
DB 1504 CTGGGCTGTGACAGAAACCCCGTTACGCGGCAACGCTGCGCTTATCCGGAATCAT 1445  
OY 2219 CGTCTTGAATCCAAACCGGTAAAGACAGACTTATCCCACTGGCAGAGCACTGTATAC 2278  
DB 1444 CGTCTTGAATCCAAACCGGTAAAGACAGACTTATCCCACTGGCAGAGCACTGTATAC 1385  
OY 2279 AGGATTAGCAGAGGAGATATGAGCGGTCTACAGAGTTCTTGAAGTGTGCGCTAAC 2338

Db 1384 AGAATTAGCAGACCGAGATGATGAGCGGATGCTACAGATTCTTAAGTGGTGGCTTAAC 1325  
 Qy 2339 TACGGCTACCTAGAAAGA CAGTATTTGGTATCTGCGCTGCTGGAAGCCGTTACCTTC 2398  
 Db 1324 TACGGCTACCTAGAAAGACAGTATTTGGTATCTGCGCTGCTGGAAGCCGTTACCTTC 1265  
 Qy 2399 GGAAGAAAGTGGTATGCTGTTATCCGCAACAAACACCGCTGTAGCGTGTGTTT 2458  
 Db 1264 GGAAGAAAGTGGTATGCTGTTATCCGCAACAAACACCGCTGTAGCGTGTGTTT 1205  
 Qy 2459 TTTGTTTGGACAGACAGATTACCGCGAAGAAAAAGGATCTCAAGAAAGTCTTTGATC 2518  
 Db 1204 TTTGTTTGGACAGACAGATTACCGCGAAGAAAAAGGATCTCAAGAAAGTCTTTGATC 1145  
 Qy 2519 TTTTCTACGGGGTCTGAGCTCATGTGAAAGAAAACTCACGTTAAGGATTTTGTGATG 2578  
 Db 1144 TTTTCTACGGGGTCTGAGCTCATGTGAAAGAAAACTCACGTTAAGGATTTTGTGATG 1085  
 Qy 2579 AGATTATCAAAAAGGATCTTCACTAGATCCTTTTAATTAATAAGAAAGTTTAAATCA 2638  
 Db 1084 AGATTATCAAAAAGGATCTTCACTAGATCCTTTTAATTAATAAGAAAGTTTAAATCA 1025  
 Qy 2639 ATCTAAAGTATATATAGTAACTTGGTCTGACAGTTACCAATGCTTAATCAGTAGGCA 2698  
 Db 1024 ATCTAAAGTATATATAGTAACTTGGTCTGACAGTTACCAATGCTTAATCAGTAGGCA 965  
 Qy 2699 CCTATCTCAGCGATCTGTCTATTTCTGTCATCCATAGTTCCTGACTCCCGCTGATGAG 2758  
 Db 964 CCTATCTCAGCGATCTGTCTATTTCTGTCATCCATAGTTCCTGACTCCCGCTGATGAG 905  
 Qy 2759 ATAACTACGATAGGGAGGGCTTAACATTTGGCCCGAGTGTGCAATGATACCGGAGAC 2818  
 Db 904 ATAACTACGATAGGGAGGGCTTAACATTTGGCCCGAGTGTGCAATGATACCGGAGAC 845  
 Qy 2819 CCACGCTCACCGGCTCCAGATTATACAGATTAACCAAGCCGCGAGAGGGCCGAGCGC 2878  
 Db 844 CCACGCTCACCGGCTCCAGATTATACAGATTAACCAAGCCGCGAGAGGGCCGAGCGC 785  
 Qy 2879 AGAAGTGTCTGCAACTTATCCGCTCCATCCAGTCTATTAATGTTGCGGGAAGCT 2938  
 Db 784 AGAAGTGTCTGCAACTTATCCGCTCCATCCAGTCTATTAATGTTGCGGGAAGCT 725  
 Qy 2939 AAGAGTATAGTTCGCAATTATAGTTTGCGCAACGTTGCGCATTCGTAAGGCGATC 2998  
 Db 724 AAGAGTATAGTTCGCAATTATAGTTTGCGCAACGTTGCGCATTCGTAAGGCGATC 665  
 Qy 2999 GTGATGTCAACGCTGCTGTTGATGCTTCAATTCAGCTCCGATCCCAACGATCAAG 3058  
 Db 664 GTGATGTCAACGCTGCTGTTGATGCTTCAATTCAGCTCCGATCCCAACGATCAAG 605  
 Qy 3059 CGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGTTAGCTCTTCGGTCTCCGATC 3118  
 Db 604 CGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGTTAGCTCTTCGGTCTCCGATC 545  
 Qy 3119 GTTGCGAAGTATGTTGGCGCGAGTGTATCACTAGTATGAGCAGCATGATAT 3176  
 Db 544 GTTGCGAAGTATGTTGGCGCGAGTGTATCACTAGTATGAGCAGCATGATAT 485  
 Qy 3179 TCTTCTAATGATGATCCGTAAGATGTTTTGCTGACGTTGAGTACTCAACCAAG 3238  
 Db 484 TCTTCTAATGATGATCCGTAAGATGTTTTGCTGACGTTGAGTACTCAACCAAG 425  
 Qy 3239 TCATTCGAGAATAGTATGCGGCAACGAGTGTCTTGGCCGCGCTCAATCCGGAT 3298  
 Db 424 TCATTCGAGAATAGTATGCGGCAACGAGTGTCTTGGCCGCGCTCAATCCGGAT 365  
 Qy 3299 AATACCGCGCCATAGCAGAACTTTAAAGGCTCATCATTTGAAAAGCTTCTTCCGGG 3358  
 Db 364 AATACCGCGCCATAGCAGAACTTTAAAGGCTCATCATTTGAAAAGCTTCTTCCGGG 305  
 Qy 3359 CGAAAACCTCAGAGATCTTACCGCTGTTGAGATCCAGTTGATGTAACCACTCCGATCA 3418

Db 304 CGAAAACCTCAGAGATCTTACCGCTGTGAGATCCAGTTGATGTAACCACTCGTGA 245  
 Qy 3419 CCCAATCTATTTGAGCATCTTTTACCTTTACCAAGCTTTCTGGTGAAGCAAAAAGGA 3478  
 Db 244 CCCAATCTATTTGAGCATCTTTTACCTTTACCAAGCTTTCTGGTGAAGCAAAAAGGA 185  
 Qy 3479 AGGCAAAATGCGGCAAAAAGGGAATPAGGGGCAACAGAAATGTTGAATCTCATCTC 3538  
 Db 184 AGGCAAAATGCGGCAAAAAGGGAATPAGGGGCAACAGAAATGTTGAATCTCATCTC 125  
 Qy 3539 TTTCTTTTCAATTTATTTGAGCATTTATCAGGGTTATTTGTCTCATGAGCGGATACATA 3598  
 Db 124 TTTCTTTTCAATTTATTTGAGCATTTATCAGGGTTATTTGTCTCATGAGCGGATACATA 65  
 Qy 3599 TTTGAATGATTTTGAAGAAAAATTAACAAATPAGGGGTTCCGCGCAATTTCCGAAAAGTG 3658  
 Db 64 TTTGAATGATTTTGAAGAAAAATTAACAAATPAGGGGTTCCGCGCAATTTCCGAAAAGTG 5  
 Qy 3659 CCAC 3662  
 Db 4 CCAC 1

RESULT 8  
 ADM68462/c  
 ID ADM68462 standard; DNA; 3637 BP.  
 AC ADM68462;  
 XX  
 DT 03-JUN-2004 (first entry)  
 XX  
 DS Wild type green fluorescent protein, GFP, gene construct.  
 XX  
 KW ds; mismatch endonuclease; endonuclease; gene shuffling technology;  
 KW single nucleotide polymorphism; cancer susceptibility;  
 KW sequence variation redistribution; GFP; green fluorescent protein; gene.  
 OS Aequorea victoria.  
 OS Synthetic.  
 XX  
 PN US2003157682-A1.  
 XX  
 PD 21-AUG-2003.  
 XX  
 PP 31-JAN-2003; 2003US-00356708.  
 XX  
 PR 01-FEB-2002; 2002US-0353722P.  
 PR 14-MAR-2002; 2002US-00098155.  
 PR 01-AUG-2002; 2002US-00211079.  
 XX  
 PA (PADG/) PADGETT H S.  
 PA (VAEW/) VAENHONGS A A.  
 PA (VOUT/) VOUTANT F S.  
 PA (SMIT/) SMITH M L.  
 PA (LIND/) LINDBO J A.  
 PA (FITZ/) FITZMAURICE W P.  
 PI Padgett HS, Vaenhongs AA, Vojdani FS, Smith ML, Lindbo JA;  
 PI Fitzmaurice WP;  
 DR WPI; 2003-766176/72.  
 XX  
 PT Making a mismatch endonuclease, useful in gene shuffling and in detection  
 PT of single nucleotide polymorphisms, comprises transfecting a host with a  
 PT recombinant viral vector including a polynucleotide encoding a mismatch  
 PT endonuclease.  
 XX  
 PS Example 5; SEQ ID NO 31; 79pp; English.  
 XX  
 CC The invention relates to a method of making a mismatch endonuclease  
 CC enzyme comprising transfecting a host plant, animal, yeast, fungus or  
 CC bacterium with a recombinant viral vector that encodes a polynucleotide  
 CC sequence for a mismatch endonuclease, growing the host so that the

CC polynucleotide is expressed, and extracting the mismatch endonuclease  
CC enzyme from the host. The method is useful for making mismatch  
CC endonuclease enzymes, for obtaining peptides and polynucleotides with  
CC desired functional properties and for detecting mutations. The mismatch  
CC endonuclease enzymes are useful in gene shuffling technology for  
CC developing new genes, in detecting single nucleotide polymorphisms for  
CC e.g. detecting evidence of cancer susceptibility, or in redistributing  
CC sequence variations between non-identical polynucleotide sequences. The  
CC present sequence represents the wild type green fluorescent protein, GFP,  
CC gene construct.

Sequence 3637 BP; 959 A; 879 C; 868 G; 931 T; 0 U; 0 Other;

Query Match 72.5%; Score 2656.4; DB 11; Length 3637;

Best Local Similarity 84.0%; Pred. No. 0;

Matches 3079; Conservative 0; Mismatches 556; Indels 29; Gaps 6;

```
OY 1 CTAATTGTAAAGCGTTAATATTTGTAAATTTCCGCTTAATTTTGTAAATCAGCTC 60
DB 3637 CTAATTGTAAAGCGTTAATATTTGTAAATTTCCGCTTAATTTTGTAAATCAGCTC 3578
OY 61 ATTTTAAACCAATAGCCGCAAAATCCGCAAAATCCCTTAATAATCAAAAGATAGACGA 120
DB 3577 ATTTTAAACCAATAGCCGCAAAATCCGCAAAATCCCTTAATAATCAAAAGATAGACGA 3518
OY 121 GATAGGTTGAGTGTGTTCCAGTTTGGAACAAGTCCACTATTAAAGAGCTGACTC 180
DB 3517 GATAGGTTGAGTGTGTTCCAGTTTGGAACAAGTCCACTATTAAAGAGAGTGTGACTC 3458
OY 181 CAACGTCAAAGGCGCAAAACCGTCTATCAGGGCGATGGCCCACTACGTAACCATCAC 240
DB 3457 CAACGTCAAAGGCGCAAAACCGTCTATCAGGGCGATGGCCCACTACGTAACCATCAC 3398
OY 241 CTAATCAAGTTTTTGGGGTCGAGGTGCGTTAAGCACTAAATCGAAACCTTAAAGGAG 300
DB 3397 CTAATCAAGTTTTTGGGGTCGAGGTGCGTTAAGCACTAAATCGAAACCTTAAAGGAG 3338
OY 301 CCCCCGATTTAGAGCTTTGACCGGGAAAGCCCGCAACGTCGCGAAGAAAGAGGAGAA 360
DB 3337 CCCCCGATTTAGAGCTTTGACCGGGAAAGCCCGCAACGTCGCGAAGAAAGAGGAGAA 3278
OY 361 AGCGAAAGAGCGGGCGCTGAGGGCGCTGCAAGGTAGCGGTCACTGCGCGTAACAC 420
DB 3277 AGCGAAAGAGCGGGCGCTGAGGGCGCTGCAAGGTAGCGGTCACTGCGGTACACAC 3218
OY 421 CACACCCGCGCGCTTAAATGCGCGCTACAGGGCGCGTCCCAATTCGCCATTGAGGCTCG 480
DB 3217 CACACCCGCGCGCTTAAATGCGCGCTACAGGGCGCGTCCCAATTCGCCATTGAGGCTCG 3158
OY 481 CAACGTGTGGGAAGGGCGATTCGTGCGGCGCTTTCGCTATTACGCCAGCTGGCGAAG 540
DB 3157 CAACGTGTGGGAAGGGCGATTCGTGCGGCGCTTTCGCTATTACGCCAGCTGGCGAAG 3098
OY 541 GGGATGTCTGCAAGCGATTAAGTTGGGTAAACGCGCAGGGTTTCCAGTACAGCGTTG 600
DB 3097 GGGATGTCTGCAAGCGATTAAGTTGGGTAAACGCGCAGGGTTTCCAGTACAGCGTTG 3038
OY 601 TAAACGACGCGCACTGACGCGCGCTTCCTTCAATTCACGTTTTTGAACCCGTGAGAGCG 660
DB 3037 TAAACGACGCGCACTGACGCGCGGTATACGACTCATATAGGCGAATTGGAGCTCA 2978
OY 661 GCAAGCTGCGCGTGAATGTGTTTACAGCCTGATGAGACAGATGAAGATGCTGCACAC 720
DB 2977 CCGCGGTGGCGCGCTCTAGAACATAGTGAATCCCGCGGCTGCGAGAAATCTTATTTGT 2918
OY 721 GCTCAGAACACGACGCTAGATTAACTCTAGAAAGATATATCATATTGAGAGTAACTTAA 780
DB 2917 ATAATGATCATGATGATGATTAATCCAGCAGCTGTTACAACTCAAGAGAGACATGT 2858
OY 781 AGATTAATC-ATGCGTAAATTTGACGATGGGATCTGTAAATACGACTCATATAGGGCA 838
DB 2857 GGTCTCTCTTTCTGTTGGGATCTTTGAAAGGCGAGATTTGTGTGACAGGTAATGTTGT 2798
```

```
OY 839 ATTGCGTACCGGGCCCCCTCGAGTCGACGTAATCATATGCTGTATATCGAATTCCT 898
DB 2797 CTGTAAAGAGACAGGGCGATCGCAATTTGAGATATTTGTGATATGCTGTAGTT 2738
OY 899 GCAGCCCGGGGATATCACTAGTTCTAGACGGCGGCCACCGCGGTGAGCTTCAGCTTTT 958
DB 2737 GAGCGCTTCATCTTCAATGTTGTGTCTAAATTTGAAGTTAACTTTGATTCATCTTTT 2678
OY 959 GTTCCCTTAAAGAGGTTAATTAGATCCATGCGGTCAATTTTACGACACTATCTTCT 1018
DB 2677 GTTGTCTGCGCATATGATATGATGTGTGA-GTTATATGTTATTCAAATTTGTGTCA 2619
OY 1019 AGGTTATCTAGCTGATCAGATCATATGTGTGGGCTTTTTCGGCTCAGTATCG 1078
DB 2618 AGAATGTTTCATCTTTTAAATCAATACCTTTTAACTGATTC-----TATTAAACA 2564
OY 1079 CCCAAGCTGGCGCTATCTGGGCGATCGGGGAGAAAGACCCGTCCTTTCCCGGAGGT 1138
DB 2563 GGGTATCATCTTCAAACTTGACTTC-----AGCAGTGTCTTGTAGTTCCGTC 2515
OY 1139 TGAAGCGCATGAAAGAGTTTGCAGAGATGACTGCTGCATTTGACGTTAGCGAAA 1198
DB 2514 ATCTTGAATAATATATGTTCTTTCCTGTACATTAACCTTCGGGCACTGCACTTGAATA 2455
OY 1199 ACGACCTTTACATGATGATTCGGGAAAGTGTGGCCATGACCGCTTTAACGTTGAAC 1258
DB 2454 GTCATGCCGTTTCATATGATCTGGGTATCTTGAAGAAAGCATTTGAACACCATTAAGAAAGT 2395
OY 1259 GTTGGTTCAGGCGCACTGGGATACAGATTGCTGCGGCTTTTCGGACACAGTTCCGGAT 1318
DB 2394 AGTACAAAGTTTGGCCATGGAACAGTATGTTTTCAGTATGTGCAATTAATTAAGGT 2335
OY 1319 GGTACGCCGAAGCGCATACAGCAACCCGAACAATCCGCGCACAGCCGGAACCTCGTGC 1378
DB 2334 AAGTTTTC-----GTATGTTGATCACTTCACTCTCTCCATGACAGAAATTTGTGC 2280
OY 1379 CCGGTGCAATTAATGACAGCGGTGGCGCTGGGATTTACGTACGACGAGCGGTA 1438
DB 2279 CCATTAACATGACCATCTAATTAACAAGAAATTTGGGCAACTCCAGTAAAGATTTCTTCT 2220
OY 1439 TCCTGGCTGATGTCGCGAGAAATGACATGGAATACCCCGTGAATTAACCGGCGGCGCGC 1498
DB 2219 CTTTATCTCATCGGTACCCAGCTTTTGTCTCTTAAATGAGAGTTA-----ATTGCGGC 2165
OY 1499 TTGCGTAAATCATGATCATAGCTGTTTCTGTGTGAATTTGTTATTCGCTCACAAATTC 1558
DB 2164 TTGGCGTAAATCATGATCATAGCTGTTTCTGTGTGAATTTGTTATTCGCTCACAAATTC 2105
OY 1559 CACAACTATGAGCCGGAAGATTAAGTTAAAGCTGGGGTGCCTTAATGATGAGCTTA 1618
DB 2104 CACAACTATGAGCCGGAAGATTAAGTTAAAGCTGGGGTGCCTTAATGATGAGCTTA 2045
OY 1619 CTCACATTAATTTGGTTGCGCTGCTGCTGCGCTTCCAGTGGGAAACCTTTCGTCGACG 1678
DB 2044 CTCACATTAATTTGGTTGCGCTGCTGCTGCGCTTCCAGTGGGAAACCTTTCGTCGACG 1965
OY 1679 CTGCATTAATGAAATCGGCCAAACGCGCGGGAGAGCGGTTTTCGTTATTTGGCGCTTTC 1738
DB 1964 CTGCATTAATGAAATCGGCCAAACGCGCGGGAGAGCGGTTTTCGTTATTTGGCGCTTTC 1925
OY 1739 GCTTCTCTGCTCATCTGACTGCTGCGCTCGGTGCTTGGCGGTGGCGGAGCGGTAATCAGCT 1798
DB 1924 GCTTCTCTGCTCATCTGACTGCTGCGCTCGGTGCTTGGCGGTGGCGGAGCGGTAATCAGCT 1865
OY 1799 CACTCAAGGCGGTAAATACGGTTATTCACAGAAATCAGGGGTAACGAGGAGAAACATG 1858
DB 1864 CACTCAAGGCGGTAAATACGGTTATTCACAGAAATCAGGGGTAACGAGGAGAAACATG 1805
OY 1859 TGAGCAAAAGGCGCAAAAGGCGAGAACCTGTAATAAGGCGCGGTTGCTGGCTTTTTC 1918
DB 1804 TGAGCAAAAGGCGCAAAAGGCGAGAACCTGTAATAAGGCGCGGTTGCTGGCTTTTTC 1745
OY 1919 CATAGGCTCCGCCCCCTGACAGACATCAAAAAATGACGCTCAAGTCAAGGTGGCGA 1978
```



Db 1744 CATAGGCTCCGCCCCCTGACGAGCATACACAAAAATCAGCCTAAGTCAAGAGGTGGCGA 1685  
Qy 1979 AACCCGACAGGACTATAAAGATACAGGCGTTCCTCCCTGGAAGCTCCCTCGGCTCT 2038  
Db 1684 AACCCGACAGGACTATAAAGATACAGGCGTTCCTCCCTGGAAGCTCCCTCGGCTCT 1625  
Qy 2039 CCTGTTCGACCCCTCGGCTTACCGGAACTCTGTCGCCCTTCTCCCTTCGGGAAGCGTG 2098  
Db 1624 CCGTTCGACCCCTCGGCTTACCGGAACTCTGTCGCCCTTCTCCCTTCGGGAAGCGTG 1565  
Qy 2099 GCGCTTTCATACGTCACGCTGAGATCTCAGTTCGGGTGAGGCTGTCCTCAAG 2158  
Db 1564 GCGCTTTCATACGTCACGCTGAGATCTCAGTTCGGGTGAGGCTGTCCTCAAG 1505  
Qy 2159 CTGGCTGTGTGACAGAAACCCCGTTCAGCCGACCGCTGCGCTTATCCGGTAACTAT 2218  
Db 1504 CTGGCTGTGTGACAGAAACCCCGTTCAGCCGACCGCTGCGCTTATCCGGTAACTAT 1445  
Qy 2219 CGTCTTGAATCAACCCGCTAAGACAGACTATACGACTGGCAGACAGCACTGGTAACT 2278  
Db 1444 CGTCTTGAATCAACCCGCTAAGACAGACTATACGACTGGCAGACAGCACTGGTAACT 1385  
Qy 2279 AGGATTAGCAGAGGAGGTATAGGCGGTCTACAGAGTTCCTTGAAGGTGGCTTAC 2338  
Db 1384 AGGATTAGCAGAGGAGGTATAGGCGGTCTACAGAGTTCCTTGAAGGTGGCTTAC 1325  
Qy 2339 TACGGCTACACTAGAAAGACAGTATTTGGTATCTGCGCTGCTGAAAGCGCTTATC 2398  
Db 1324 TACGGCTACACTAGAAAGACAGTATTTGGTATCTGCGCTGCTGAAAGCGCTTATC 1265  
Qy 2399 GGAAGAAAGAGTGGTGTCTTGAATCCGGCAAAACAACCGCTGTGTAGCGGTGTTT 2458  
Db 1264 GGAAGAAAGAGTGGTGTCTTGAATCCGGCAAAACAACCGCTGTGTAGCGGTGTTT 1205  
Qy 2459 TTTGTTTGAACAGCAGATTACGGCGAGAAAAAAGATCTCAAGAAAGATCCTTTGATC 2518  
Db 1204 TTTGTTTGAACAGCAGATTACGGCGAGAAAAAAGATCTCAAGAAAGATCCTTTGATC 1145  
Qy 2519 TTTTTCACGGGGTCTGACGCTCAGTGAACGAAAATCTCAGTTAAGGATTTTGGTATG 2578  
Db 1144 TTTTTCACGGGGTCTGACGCTCAGTGAACGAAAATCTCAGTTAAGGATTTTGGTATG 1085  
Qy 2579 AATTTATCAAAAAGAGTCTTCACTAGATCTTTTAAATTTAAATGAAGTTTAAATCA 2638  
Db 1084 AATTTATCAAAAAGAGTCTTCACTAGATCTTTTAAATTTAAATGAAGTTTAAATCA 1025  
Qy 2639 ATCTAAGTATATATGAGTAACTTGGTCTGACAGTTACCAATGCTTAACTAGTAGGCA 2698  
Db 1024 ATCTAAGTATATATGAGTAACTTGGTCTGACAGTTACCAATGCTTAACTAGTAGGCA 965  
Qy 2699 CCTATCTCAGGATCTGTCTATTTCTGTCATCCATAGTGGCTGATCCCGCTGTGTAG 2758  
Db 964 CCTATCTCAGGATCTGTCTATTTCTGTCATCCATAGTGGCTGATCCCGCTGTGTAG 905  
Qy 2759 ATTACTACGATACGGGAGGGCTTACATCTGCGCCCAAGTGTGCAATGATACCGAGAC 2818  
Db 904 ATTACTACGATACGGGAGGGCTTACATCTGCGCCCAAGTGTGCAATGATACCGAGAC 845  
Qy 2819 CCACGCTCACCCTCCAGATTATACGAATAAACAACACCGGAGGGCCGAGCGC 2878  
Db 844 CCACGCTCACCCTCCAGATTATACGAATAAACAACACCGGAGGGCCGAGCGC 785  
Qy 2879 AGAAGGTGCTCCCAACTTTATCCGCTCATCCAGTCTATTAATTTGTTGCGGGAAGCT 2938  
Db 784 AGAAGGTGCTCCCAACTTTATCCGCTCATCCAGTCTATTAATTTGTTGCGGGAAGCT 725  
Qy 2939 AGAGTAAGTATGCGCAGTTAATAGTTGCGCAACGTTTGGCATTTGCTACAGGATC 2998  
Db 724 AGAGTAAGTATGCGCAGTTAATAGTTTGGCAACGTTTGGCATTTGCTACAGGATC 665  
Qy 2999 GTGGTTCACGCTCGTGTGTGTATGCTTCAATCACTCCGGTTCACCAAGCATCAAG 3058  
|||||

Db 664 GTGGTTCACGCTCGTGTGTATGATGGCTTCAATCACTCCGGTTCACCAAGCATCAAG 605  
Qy 3059 CGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGTTAGCTCTTCGCTCCGATC 3118  
Db 604 CGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGTTAGCTCTTCGCTCCGATC 545  
Qy 3119 GTTGTCAAGATTAAGTTGGCCGAGTTATCACTCATAGTTATGACAGATCACTAAT 3178  
Db 544 GTTGTCAAGATTAAGTTGGCCGAGTTATCACTCATAGTTATGACAGATCACTAAT 485  
Qy 3179 TCTCTTACTGATGACCATCCGTAAGATGTTTTCTGTGACTGTGATCACTCAACAG 3238  
Db 484 TCTCTTACTGATGACCATCCGTAAGATGTTTTCTGTGACTGTGATCACTCAACAG 425  
Qy 3239 TCATTTCAAGATTAAGTTATGCGGCGACCGAGTTGCTTGTGCGGCGCAATACGGGAT 3298  
Db 424 TCATTTCAAGATTAAGTTATGCGGCGACCGAGTTGCTTGTGCGGCGCAATACGGGAT 365  
Qy 3299 AATACCGGCGCACATAGCAGAACTTTAAAGTCTCATCATTTGAAAAAGTTTCTGCGG 3358  
Db 364 AATACCGGCGCACATAGCAGAACTTTAAAGTCTCATCATTTGAAAAAGTTTCTGCGG 305  
Qy 3359 CGAAAACTCTCAAGATCTTACCGCTGTGTAGATCAGTTGATGATACCACTCGGCA 3418  
Db 304 CGAAAACTCTCAAGATCTTACCGCTGTGTAGATCAGTTGATGATACCACTCGGCA 245  
Qy 3419 CCCAATGATCTTCAAGATCTTTTACCTTTACCAAGCTTTCTGGGTGAGCAAAAAACAGA 3478  
Db 244 CCCAATGATCTTCAAGATCTTTTACCTTTACCAAGCTTTCTGGGTGAGCAAAAAACAGA 185  
Qy 3479 AGGCAAAATGCGCAAAAAAGGAATAGGCGCAGACAGAAATGTGAATATCATCTC 3538  
Db 184 AGGCAAAATGCGCGCAAAAAAGGAATAGGCGCAGACAGAAATGTGAATATCATCTC 125  
Qy 3539 TTCCCTTTTCAATTAATTAAGCAATTAATCAGGCTTATGTCATGACGATACATA 3598  
Db 124 TTCCCTTTTCAATTAATTAAGCAATTAATCAGGCTTATGTCATGACGATACATA 65  
Qy 3599 TTTGATGATTAATTAAGAAAAATTAACAATAGGGGTCCGCGCATTTCCCGCAAAAGTG 3658  
Db 64 TTTGATGATTAATTAAGAAAAATTAACAATAGGGGTCCGCGCATTTCCCGCAAAAGTG 5  
Qy 3659 CCAC 3662  
Db 4 CCAC 1  
RESULT 9  
ADP2615/c  
ID ADP2615 standard; DNA; 3637 BP.  
AC ADP2615;  
DT 26-AUG-2004 (first entry)  
XX Green fluorescent protein (GFP) plasmid DNA #1.  
XX Sequence variation; heteroduplex; transcription; DNA integration;  
XX ribozyme expression; gene; ds; green fluorescent protein; GFP.  
OS Aequorea victoria.  
XX Synthetic.  
XX US2004110130-A1.  
PD 10-JUN-2004.  
XX 25-OCT-2002; 2002US-00280913.  
PF 02-FEB-2001; 2001US-0266386P.  
PR 14-FEB-2001; 2001US-0268785P.  
PR 01-FEB-2002; 2002US-0006390.  
PR 08-AUG-2002; 2002US-0402342P.



XX (LARG-) LARGE SCALE BIOLOGY CORP.  
XX  
XX Padgett HS, Lindbo JA, Fitzmaurice WP;  
XX WPI; 2004-440326/41.  
XX  
PT Redistributing sequence variations between non-identical polynucleotide  
PT sequences, useful for generating improved polynucleotide having a desired  
PT characteristic, comprises making a heteroduplex and introducing a nick.  
XX  
XX Example 5; SEQ ID NO 31; 75bp; English.  
XX  
XX The invention relates to an in vitro method of redistributing sequence  
XX variations between non-identical polynucleotide sequences, comprising  
XX making a heteroduplex polynucleotide from two non-identical  
XX polynucleotides, introducing a nick in the second strand at or near a  
XX base pair mismatch site, removing the mismatched base(s) from the  
XX mismatch site where the nick occurred and using the first strand as a  
XX template to replace the removed base(s) with bases that complement the  
XX base(s) in the first strand. The invention also relates to an in vitro  
XX method of making a population of sequence variants from a heteroduplex  
XX polynucleotide sequence, obtaining a polynucleotide sequence encoding a  
XX desired functional property and identifying a reinserted DNA molecule  
XX encoding a protein with a desired functional property. The method is  
XX useful for generating an improved polynucleotide sequence or a population  
XX of improved polynucleotide sequences possessing at least one desired  
XX phenotypic characteristic (e.g., promotes transcription of linked  
XX polynucleotides), where such polynucleotides are useful for expression  
XX from a plant, animal, fungal, yeast, or bacterial expression vector, for  
XX integration to form a transgenic plant, animal or microorganism, and for  
XX expression of a ribozyme. This sequence represents DNA used in the scope  
XX of the invention.  
SQ Sequence 3637 BP; 959 A; 879 C; 868 G; 931 T; 0 U; 0 Other;  
Query Match 72.5%; Score 2656.4; DB 12; Length 3637;  
Best Local Similarity 84.0%; Pred. No. 0;  
Matches 3079; Conservative 0; Mismatches 556; Indels 29; Gaps 6;  
QY 1 CTAAATTGTAAGCGTTAATATTTTGTAAATTCGCGTTAAATTTTGTAAATCAGCTC 60  
DB 3637 CTAAATTGTAAGCGTTAATATTTTGTAAATTCGCGTTAAATTTTGTAAATCAGCTC 3578  
QY 61 ATTTTAAACCAATAGGCGCAATCGCAAAATCCCTTAATATCAAAAGAAATAGCCGA 120  
DB 3577 ATTTTAAACCAATAGGCGCAATCGCAAAATCCCTTAATATCAAAAGAAATAGCCGA 3518  
QY 121 GATAGGGTGAAGTGTTCAGTTTGGAAACAAGTCCACTATTAAGAAACGTGACTC 180  
DB 3517 GATAGGGTGAAGTGTTCAGTTTGGAAACAAGTCCACTATTAAGAAACGTGACTC 3458  
QY 181 CAACGTCAAAAGGCGCAAAACCGTCTATCAGGGCGATGCGCCACTAGCTGAACCATCAC 240  
DB 3457 CAACGTCAAAAGGCGCAAAACCGTCTATCAGGGCGATGCGCCACTAGCTGAACCATCAC 3398  
QY 241 CTAAATCAAGTTTTTGGGGTCTGAGGTGCGCTAAAGCACTAAATCGAAACCTTAAGGAG 300  
DB 3397 CTAAATCAAGTTTTTGGGGTCTGAGGTGCGCTAAAGCACTAAATCGAAACCTTAAGGAG 3338  
QY 301 CCCCCGATTAGAGCTTGAACGGGGAACCGGCGCAACGTGCGGAAAGAAAGGAAAGAA 360  
DB 3337 CCCCCGATTAGAGCTTGAACGGGGAACCGGCGCAACGTGCGGAAAGAAAGGAAAGAA 3278  
QY 361 AGGAAAGAGAGCGGCGCTAGAGGCGCTGCAAGGTAGCGGTACGCTGCGCTAACAC 420  
DB 3277 AGGAAAGAGAGCGGCGCTAGAGGCGCTGCAAGGTAGCGGTACGCTGCGCTAACAC 3218  
QY 421 CACACCGCGCGCTTAATGCGCGCTACAGAGGCGCTCCATTGCGCTTACAGGCTGCG 480  
DB 3217 CACACCGCGCGCTTAATGCGCGCTACAGAGGCGCTCCATTGCGCTTACAGGCTGCG 3158  
QY 481 CAACTGTTGGGAAGGGGATCGGTGCGGCGCTCTTCGCTATTAGCCACGCTGGGAAAG 540

DB 3157 CAACTGTTGGGAAGGGGATCGGTGCGGCGCTCTTCGCTATTAGCCACGCTGGGAAAG 3098  
QY 541 GGAATGCTGTCGAAGCGGATTAAGTTGGTAAACGCCAGGGTTTTCCAGTACGACGTTG 600  
DB 3097 GGAATGCTGTCGAAGCGGATTAAGTTGGTAAACGCCAGGGTTTTCCAGTACGACGTTG 3038  
QY 601 TAAACGACCGCGCTAGAGGCGGCTCGTTCATTACAGTTTAAACCCGTGAGGACGG 660  
DB 3037 TAAACGACCGCGCTAGAGGCGGCTCGTTCATTACAGTTTAAACCCGTGAGGACGG 2978  
QY 661 GCACATCGCGGTGCAATGTTTACAGGTGATGAGCAAGTGAAGATGCTGACAC 720  
DB 2977 CCGGCGTGGGCGCGCTCTGAAGTAAGTATGATCCCGGCTGACGAATTTTATTTGCT 2918  
QY 721 GCTGCAAGAACGACGAGCTAATTAACCTTAAGAAAGATTAATATTTGATGATGATTA 780  
DB 2917 ATAGTATCATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2858  
QY 781 AGATATTC--ATGCTTAAATTTGACGATGATGATGATGATGATGATGATGATGATGAT 838  
DB 2857 GGTCTCTTTTGTGGGATCTTTCGAAAGGCGAGATGATGATGATGATGATGATGATGAT 2798  
QY 839 ATTGCGTACCGGCG 898  
DB 2797 CTGCTTAAAGGACAGGCGCATGCCAATGGAATTTTGTGATTAATGCTGCTAGTT 2738  
QY 899 GCAAGCCCGGGGATATCACTAGTTTCTAGAGGCGCGCGCGCGCGCGCGCGCGCGCGCT 958  
DB 2737 GAACGCTTCAATCTTAATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 2678  
QY 959 GTTCCCTTTAGTGAAGGTTAATTAATCCATGCTGATGATTTTACGACATCTTTCT 1018  
DB 2677 GTTGTCTGCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2619  
QY 1019 AGGTTATATCATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1078  
DB 2618 AGAATGTTTCATCTTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2564  
QY 1079 CCCAAGCTGCGCTATCTGCGCATCGGGAAGAAAGCCCGCTTTTCCGCGAGGT 1138  
DB 2563 GGGATACCTTCAAACTTACCTTC-----AGCAGGTCTTGTAGTTCCGCTC 2515  
QY 1139 TGAAGCGCATGGAAGAGTTTGGCGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1198  
DB 2514 ATCTTGAAAAATATATGTTCTTCTGCTGATTAACCTTCGCGCATGCTGCTGAAAAA 2455  
QY 1199 ACGCAGTTTACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1258  
DB 2454 GTCAATGCTTTTCAATATGATCTGGGATCTTGAAGAAATTAACCATTAAGAAAGT 2395  
QY 1259 GTTCTGTTCAAGCACTGAGATACAGTTTCTGCGGCTTTTCCGACACAGTTTCCGAT 1318  
DB 2394 AGTGAACAAGTTGGCCATGGAACAGTATTTTCCAGTATGTCGAATAAATTAAGGT 2335  
QY 1319 GGTACGCGGAAGGATCAAGCAACCCGAACAATACCGGCAACGCGGAATGCGGTG 1378  
DB 2334 AAGTTTTC-----GTAATGTTGATCACTTCACCTCTCAGTACAGAAAAATTTGTG 2280  
QY 1379 CGATGTCGATTAATGAACAGCGGTGCGGCTGGGATTAATGCTACGACGAGAGCGGTA 1438  
DB 2279 CCAATTAACATCACATCTTAATTAACAAGAAATGGGCAACTCAAGTAAAGTTCTTCT 2220  
QY 1439 TCCTGCTGATGCTCCGCAAAATGACATGATACCCCGTGAATGTTACCCGCGCGCGCG 1498  
DB 2219 CTTTACTATATGCTGATCCAGCTTTTGTTCCTTTAGTGAAGGTTA-----ATTGCGCG 2165  
QY 1499 TTGCGCTAATCATGCTGATAGCTGTTTCTGTGTGAATTTTATTCGCTCACAATTTCA 1558  
DB 2164 TTGCGCTAATCATGCTGATAGCTGTTTCTGTGTGAATTTTATTCGCTCACAATTTCA 2105  
QY 1559 CACAACATACAGCGGGAACATTAAGTGAAGCGTGGGTCCTTAATGAGGAGTTAA 1618

Db 2104 CACAACATAGAGCCGGAACATTAAGTAAAGCTGGGATGCTTAATGAGTAGCTAA 2045  
Qy 1619 CTACATTAATTTGGGTTGGCTTCACTGCCGCTTTCAGTGGGAAACCTTCCGCGAG 1678  
Db 2044 CTACATTAATTTGGGTTGGCTTCACTGCCGCTTTCAGTGGGAAACCTTCCGCGAG 1985  
Qy 1679 CTGCATTAATTAATTCGGCCAAAGCGGGGAGAGCGGTTTGGCTATTTGGGCGCTTTC 1738  
Db 1984 CTGCATTAATTAATTCGGCCAAAGCGGGGAGAGCGGTTTGGCTATTTGGGCGCTTTC 1925  
Qy 1739 GCTTCCTGGCTCACTGACTGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 1798  
Db 1924 GCTTCCTGGCTCACTGACTGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 1865  
Qy 1799 CACTCAAAAGCGGTAATACGGTTATCCAGAAATCAAGGAAATACCGAGAAAGAAATG 1858  
Db 1864 CACTCAAAAGCGGTAATACGGTTATCCAGAAATCAAGGAAATACCGAGAAAGAAATG 1805  
Qy 1859 TGAGCAAAAGCGGTAATACGGTTATCCAGAAATCAAGGAAATACCGAGAAAGAAATG 1918  
Db 1804 TGAGCAAAAGCGGTAATACGGTTATCCAGAAATCAAGGAAATACCGAGAAAGAAATG 1745  
Qy 1919 CATAGGCTCCGCCCCCTGACGAGCAATCAAAATTCAGCTCAAGTCAAGTGGGCA 1978  
Db 1744 CATAGGCTCCGCCCCCTGACGAGCAATCAAAATTCAGCTCAAGTCAAGTGGGCA 1665  
Qy 1979 AACCGACAGAGCTATTAAGATACAGAGGCTTCCCTGGAAGGCTCCCTGGTGGCTCT 2038  
Db 1684 AACCGACAGAGCTATTAAGATACAGAGGCTTCCCTGGAAGGCTCCCTGGTGGCTCT 1625  
Qy 2039 CTTGTTCCGACCTGCGCTTACCGGATACCTGCTCCCTTCTCCCTGGGAAAGCGTG 2098  
Db 1624 CTTGTTCCGACCTGCGCTTACCGGATACCTGCTCCCTTCTCCCTGGGAAAGCGTG 1565  
Qy 2099 GCGCTTTCATAGCTTCAAGCTGTAAGTATCTCAAGTGGTGGTGGTGGTGGTGGTGGTGG 2158  
Db 1564 GCGCTTTCATAGCTTCAAGCTGTAAGTATCTCAAGTGGTGGTGGTGGTGGTGGTGGTGG 1505  
Qy 2159 CTGGGCTGTGTGACGAACCCCGCTGAGCCCAACCGCTGCGCTTATCCGTAATCTAT 2218  
Db 1504 CTGGGCTGTGTGACGAACCCCGCTGAGCCCAACCGCTGCGCTTATCCGTAATCTAT 1445  
Qy 2219 CGTCTTGAAGTCCACCGGTAAAGACAGACTTATGCGCACTGGGACAGCACTGGTAC 2278  
Db 1444 CGTCTTGAAGTCCACCGGTAAAGACAGACTTATGCGCACTGGGACAGCACTGGTAC 1385  
Qy 2279 AGGATTAGCAGAGCGAGTATGTAGGCGGTCTACAGATTTCTTGAAGTGGGCTTAAC 2338  
Db 1384 AGGATTAGCAGAGCGAGTATGTAGGCGGTCTACAGATTTCTTGAAGTGGGCTTAAC 1325  
Qy 2339 TACGGCTACATTAAGAGACATTTTGTATCTGCGCTGCTGTAAGGCACTTACCTTC 2398  
Db 1324 TACGGCTACATTAAGAGAGACATTTTGTATCTGCGCTGCTGTAAGGCACTTACCTTC 1265  
Qy 2399 GGAAGAAAGATGTGATCTTGTATCCGCAAAACCAACCGCTGGTGGTGGTGGTGGTGGTGG 2458  
Db 1264 GGAAGAAAGATGTGATCTTGTATCCGCAAAACCAACCGCTGGTGGTGGTGGTGGTGGTGG 1205  
Qy 2459 TTTGTTTGGCAAGAGATTTAGCGGCAAAAGATCTCAAGAGATCTTTTATC 2518  
Db 1204 TTTGTTTGGCAAGAGATTTAGCGGCAAAAGATCTCAAGAGATCTTTTATC 1145  
Qy 2519 TTTTCTTACGCGGCTCTGACGCTGAGTGAACGAAATCTACGTTAAGGATTTTGGTATG 2578  
Db 1144 TTTTCTTACGCGGCTCTGACGCTGAGTGAACGAAATCTACGTTAAGGATTTTGGTATG 1085  
Qy 2579 AGATTATCAAAAGATCTTCACTGATCTTTTAAATTAATAAGATTTTAAATCA 2638  
Db 1084 AGATTATCAAAAGATCTTCACTGATCTTTTAAATTAATAAGATTTTAAATCA 1025  
Qy 2639 ATCTAAGTATATAGTAAACTTGGTCTGACGTTAACAAGCTTATCACTGAGGCA 2698  
Db 1024 ATCTAAGTATATAGTAAACTTGGTCTGACGTTAACAAGCTTATCACTGAGGCA 965

Qy 2699 CCTATCTACGAGATCTGTCTATTTGTTCAATCCATAGTGGCTGAGCTCCCGGCTGGTNG 2758  
Db 964 CCTATCTACGAGATCTGTCTATTTGTTCAATCCATAGTGGCTGAGCTCCCGGCTGGTNG 905  
Qy 2759 ATTAATCAAGATCGGAGGCTTACATCTGGGCTCCAGTGGTCAATGATATACCGAGAC 2818  
Db 904 ATTAATCAAGATCGGAGGCTTACATCTGGGCTCCAGTGGTCAATGATATACCGAGAC 845  
Qy 2819 CCAGGCTACCGGCTCCAGATTTATCAAGATTAACAGCCAGCCGGAAGGCGAGGCG 2878  
Db 844 CCAGGCTACCGGCTCCAGATTTATCAAGATTAACAGCCAGCCGGAAGGCGAGGCG 785  
Qy 2879 AGAAGTGTCTGCAACTTATCCGCTCCATCCAGCTTATTAATTTGGCGGAAAGCT 2938  
Db 784 AGAAGTGTCTGCAACTTATCCGCTCCATCCAGCTTATTAATTTGGCGGAAAGCT 725  
Qy 2939 AGAGTAAGTATTCGCAAGTATATAGTTGGCAAGTTGTGCAATTTGCTACAGGATC 2998  
Db 724 AGAGTAAGTATTCGCAAGTATATAGTTGTGCAAGTTGTGCTACAGGATC 665  
Qy 2999 GTGGTGTACGCTGCTGTTGGTATGGCTTCACTTCAAGCTCCGATCAAGG 3058  
Db 664 GTGGTGTACGCTGCTGTTGGTATGGCTTCACTTCAAGCTCCGATCAAGG 605  
Qy 3059 CGAGTTACATGATCCCCATGTTGTGCAAAAGCGGTTAGCTCTTGGTCTCCGATC 3118  
Db 604 CGAGTTACATGATCCCCATGTTGTGCAAAAGCGGTTAGCTCTTGGTCTCCGATC 545  
Qy 3119 GTTGTCAAGATTAAGTTGGCGCAGTGTATCACTAGTTATGGCAGCACTGCAATAT 3178  
Db 544 GTTGTCAAGATTAAGTTGGCGCAGTGTATCACTAGTTATGGCAGCACTGCAATAT 485  
Qy 3179 TCTCTTCTGTCAGGCAATCCGTAAGTCTTCTGTAGCTGTGAGTACTCAAG 3238  
Db 484 TCTCTTCTGTCAGGCAATCCGTAAGTCTTCTGTAGCTGTGAGTACTCAAG 425  
Qy 3239 TCATTCTGAAGATAGTGTATGCGGCAAGGATGCTTGGCGGCTCAATACGGAT 3298  
Db 424 TCATTCTGAAGATAGTGTATGCGGCAAGGATGCTTGGCGGCTCAATACGGAT 365  
Qy 3299 AATACGCGGCAATAGCAACTTTAAAGTGTCACTATTTGAAAGCTTCTGGG 3358  
Db 364 AATACGCGGCAATAGCAACTTTAAAGTGTCACTATTTGAAAGCTTCTGGG 305  
Qy 3359 CGAAATCTCTCAAGATCTTACCGCTGTGAGATCCAGTTGGAATACCACTGTGCA 3418  
Db 304 CGAAATCTCTCAAGATCTTACCGCTGTGAGATCCAGTTGGAATACCACTGTGCA 245  
Qy 3419 CCCAAGTATCTTCAAGCACTTTTATCTTCAACAGGCTTCTGGGTGAGCAAAACAGGA 3478  
Db 244 CCCAAGTATCTTCAAGCACTTTTATCTTCAACAGGCTTCTGGGTGAGCAAAACAGGA 185  
Qy 3479 AGCAAAATTCGCGCAAAAGGAAATAGGCGGCAAGGAAATGTTGATATCTCATATCTC 3538  
Db 184 AGCAAAATTCGCGCAAAAGGAAATAGGCGGCAAGGAAATGTTGATATCTCATATCTC 125  
Qy 3539 TTTCTTTTCAATTTATTAAGATTTATCAAGGTTATTTGTCTCATGAGCGGATACATA 65  
Db 124 TTTCTTTTCAATTTATTAAGATTTATCAAGGTTATTTGTCTCATGAGCGGATACATA 65  
Qy 3599 TTTGAATGATTTAGAAAAATTAACAATAGGGGTTCCGCGCACTTTCCGAAAGTG 3658  
Db 64 TTTGAATGATTTAGAAAAATTAACAATAGGGGTTCCGCGCACTTTCCGAAAGTG 5  
Qy 3659 CCAC 3662  
Db 4 CCAC 1

RESULT 10  
AD126329/c  
ID AD126329 standard; DNA; 3637 BP.

XX	AD126329;
AC	22-APR-2004 (first entry)
XX	
XX	Novel endonuclease Res I-related plasmid pBSC3BP.
DE	
KM	endonuclease; molecular biology; plant propagation; phenotypic trait;
KW	hericide tolerance; heat tolerance; cold tolerance; drought; salinity;
KW	osmotic stress; pest resistance; insect; nematode; arachnid; fungal;
KM	bacterial; vital; enzyme production; secondary metabolite;
KW	male sterility; female sterility; dwarfness; early maturity; Res I;
KX	pBSC3BP, de.
OS	Unidentified.
XX	
PN	US2003148315-A1.
XX	
PD	07-AUG-2003.
XX	
PF	01-AUG-2002; 2002US-00211079.
XX	
PR	01-FEB-2002; 2002US-0353722P.
XX	
PR	14-MAR-2002; 2002US-00098155.
XX	
PA	(PADG/) PADGETT H S.
XX	(VAEW/) VAETHONGS A A.
PI	Padgett HS, Vaethongs AA;
XX	
DR	WPI; 2003-897548/82.
XX	
PT	New nucleic acid molecule encoding endonucleases, useful in molecular
XX	biology, specifically to generating populations of related nucleic acid
PT	molecules, and in plant propagation with useful phenotypic traits.
XX	
PS	Example 14; Fig 5; 46pp; English.
XX	
CC	This invention relates to a novel endonuclease (Res I) nucleic acid
CC	molecule which comprises a fully defined sequence of 899 bp given in the
CC	specification. The methods and compositions of the present invention are
CC	useful in molecular biology, and more specifically to generating
CC	populations of related nucleic acid molecules. They may also be used in
CC	plant propagation with useful phenotypic traits, such as improved
CC	tolerance to herbicides, improved tolerance to extremes of heat or cold,
CC	drought, salinity or osmotic stress, improved resistance to pests
CC	(insects, nematodes or arachnids) or diseases (fungal, bacterial or
CC	viral), production of enzymes or secondary metabolites, male or female
CC	sterility, dwarfness and early maturity. The present sequence is that of
CC	a plasmid which is related to the invention.
SQ	Sequence 3637 BP; 942 A; 900 C; 860 G; 935 T; 0 U; 0 Other;
Query Match	72.3%; Score 2647.8; DB 10; Length 3637;
Best Local Similarity	83.4%; Pred. No. 0;
Matches 3054; Conservative	0; Mismatches 582; Indels 27; Gaps 3;
OY	1 CTAAATGTGAAGCGTTAATATTGGTAATAATTCGCTTAATTTGTGAATCAGCTC 60
DB	3637 CTAATTTGTGAAGCGTTAATATTGGTAATAATTCGCTTAATTTGTGAATCAGCTC 3578
OY	61 ATTTTTAACCAATGAGCGGAATGGGAAAATCCCTTAATTAACAAGAATGACCGA 120
DB	3577 ATTTTTTAACCAATGAGCGGAATGGGAAAATCCCTTAATTAACAAGAATGACCGA 3518
OY	121 GATAGGGTGAGTGTGTTCCAGTTGGACAAGAGTCACATATTAAAGACGTGACTC 180
DB	3517 GATAGGGTGAGTGTGTTCCAGTTGGACAAGAGTCACATATTAAAGACGTGACTC 3458
OY	181 CAAGCTCAAAGGCGAAAAACCGTCTATCAGGCGCATGCCCCCTAGTAGCAATCACCC 240
DB	3457 CAAGCTCAAAGGCGAAAAACCGTCTATCAGGCGCATGCCCCCTAGTAGCAATCACCC 3398

QY	241	CTAATCAAGTTTTTTGGGGTGAAGTGCCTTAAGCACTAAATGGAACCTTAAGGAG	300
Db	3397	CTAAACAAAGTTTTTGGGGTGAAGTGCCTTAAGCACTAAATGGAACCTTAAGGAG	3338
QY	301	CCCCGATTTAGAGCTTGAAGGGGAAAGCCGGCGAAGTGCGGAGAAAGGAAAGAA	360
Db	3337	CCCCGATTTAGAGCTTGAAGGGGAAAGCCGGCGAAGTGCGGAGAAAGGAAAGAA	3278
QY	361	AGCGAAGAGAGCGGGCGCTAGGGGCGTGGCAAGTGTAGCGGTCAACGCTGACAC	420
Db	3277	AGCGAAGAGAGCGGGCGCTAGGGGCGTGGCAAGTGTAGCGGTCAACGCTGACAC	3218
QY	421	CACACCCGCGCTTAAATGCGCGCTACAGGGCGCTCCATTCGCCATTACGGCTGCG	480
Db	3217	CACACCCGCGCTTAAATGCGCGCTACAGGGCGCTCCATTCGCCATTACGGCTGCG	3158
QY	481	CAACGTTGGGAAGGGGCGATGCGTGGCGGGCGCTTCCTCGTAAATTAGCGAGCTGGGAAAG	540
Db	3157	CAACGTTGGGAAGGGGCGATGCGTGGCGGGCGCTTCCTCGTAAATTAGCGAGCTGGGAAAG	3098
QY	541	GGGATGTGCTCAAGGCGAATTAACTTGGGTTAACCCAGAGTTTTCCAGTCAACAGCTTG	600
Db	3097	GGGATGTGCTCAAGGCGAATTAACTTGGGTTAACCCAGAGTTTTCCAGTCAACAGCTTG	3038
QY	601	TAAACGACGCGCAAGTAGCGCGCGCTCGTTCATTACGTTTTTGAACCCGTGAGACG	660
Db	3037	TAAACGACGCGCAAGTAGCGCGCGCTCGTTCATTACGTTTTTGAACCCGTGAGAC	2978
QY	661	GCAGACTCGCGGTCCAAATGTGTTTTACAGGTGATGAGACAGTGAAGATGCTCGACAC	720
Db	2977	CGCGGTGGCGCGCGCTCTAGAACTAGTAGATCCCGCGGTGAGAAATTTCTAATTTGT	2918
QY	721	GCTCGAGAACACGAGCTAGATTAACTTAACTTAAGATTAATCATATTTGACGTACGTTAA	780
Db	2917	ATAGTATTCATGCAATGCAATGTGTATCCAGACGAGTTACAAATCTCAAGAAAGACATGT	2858
QY	781	AGATTAATC-ATGCGTAAAAATTGACGCAATGGAGTGTGTAAATACGACTACTATAGGGCGAA	839
Db	2857	GGTCAAGCTTTTTCGTGGGATCTTTGAAAGGGCAGATGTGTGTCAGACAGGTAATGTTGT	2798
QY	840	TTGGGTACCGGGCCCCCTCGAGGTGACGAGTATCATTAAGCTTGAATTCGAATTCCTG	899
Db	2797	CTGTGAAAGGACAGGGGCATCGCCAAATGGAGTATTTGTGTAATATGCTGTGCTAGTT	2738
QY	900	CAGCCCGGGGATCCACTAGTTCTTAGAGCGGCGCCACCGGGGTGAGACTCCAGCTTTTG	959
Db	2737	GAAAGGATTCATCTTCAATGTGTGGGGAATTTGAAAGTTAGCTTGAATTCATCTTTT	2678
QY	960	TTCCCTTATGAGAGGTTAATTTAGATCCCATGCGCTCAATTTTACGACGATATCTTCTA	1019
Db	2677	GTTTGTGCGCGTATGTATATCATTTGTGTAGTTAAAGTTGTATCGAAGTTGTGTGCGA	2618
QY	1020	GGGTTAATCTAGCTGCATCAGATCATATGTCGGGTCTTTTTTCCGGCTCAGTCATCGC	1079
Db	2617	GAAATGTTCCATCTTCTTAAATCAATACCTTTTAACGTGATAC-----GATTAACAG	2563
QY	1080	CCAAGCTGGGCTAATCTGGGGCATGCGGGAAGAAAGAACCCGTCCTTCCCGGAGATT	1139
Db	2562	GGATTCACCTTCAAACTTGATCTTAGCACCGGTCTTTAGTTCCCGTCAATCTTTGAAGA	2503
QY	1140	GAAAGCGCATGGAAGAAGTTTGGCCGAGATGACTGCTGCTCATTTGACGTTGAGCGAAAA	1199
Db	2502	TATAGTGGTTCCTGTACATTAACCTTGGGGATGAGCACTTTGAAAAAAGTCAGCGCTT	2443
QY	1200	CGCACGTTTACATGATGATTCGGGAAGGTGTGCGCATGCAAGCTTTAAACGTTGAAC	1259
Db	2442	CATATGATCCCGAATAGAGAAAAAGCATTTGAACCATGAGAGAAAGTATGTGCAAGTGT	2383
QY	1260	TTGCTTACGCGCACTGGGATACAGATTCGATCGCGCTTTTCCGGAACAGCTTCGGAGT	1319
Db	2382	TGGCCATGGAACAGTGTGTTTTTCCAGTAGTGCAAATTAATTTAATGTATAGCTTTCCCTA	2323
QY	1320	GTCAGCCGGAAGCCATACGCAACCCGAACAATATCCGGGACAGCCGGAACCTGCGTGC	1379

Db 2322 TGTAGATCACTTCACTCTTCCACTGACGAAATTTGTGCCATTAAATCAGCATC 2263  
Qy 1380 GGTGTCGAGATTAAATGACAGCGGTGCGGCGCTGGAGATATACGTACAGAGCGGTTAT 1439  
Db 2262 TAAATTCACAAAGAT-----TGGACAACTCCAGTGAAMAATTC 2224  
Qy 1440 CCTGCTGGATGCGCAGAAATGACATGATACCCCGTAGATTACCGGCGGCGCT 1499  
Db 2223 TTCTCTTTTATCATGCGTACCCAGCTTTGTGTGTATGTAGAGGTTAATTGGCGCT 2164  
Qy 1500 TGGCGTATCATGTGATAGCTGTTTCTGTGTAAATTTGTTATCCGCTACAAATTCAC 1559  
Db 2163 TGGGTATCATGTGATAGCTGTTTCTGTGTAAATTTGTTATCCGCTACAAATTCAC 2104  
Qy 1560 ACAACATACGAGCGCGGAAGATTAAGTGTAAAGCTGGGGTGCCTAATAGTAGAGTAAC 1619  
Db 2103 ACAACATACGAGCGCGGAAGATTAAGTGTAAAGCTGGGGTGCCTAATAGTAGAGTAAC 2044  
Qy 1620 TCAATTAAATTCGTTGCGCTCATCTGCCCTTTCCAGTCGGGAAACCTGTGTCAGC 1679  
Db 2043 TCAATTAAATTCGTTGCGCTCATCTGCCCTTTCCAGTCGGGAAACCTGTGTCAGC 1984  
Qy 1680 TGCATTATGAATCGGCGAAACGCGCGGAGAGCGGTTTGCATATTGGCGCTCTTCG 1739  
Db 1983 TGCATTATGAATCGGCGAAACGCGCGGAGAGCGGTTTGCATATTGGCGCGTCTTCG 1924  
Qy 1740 CTTCCTGCTGACTGACTGCTGCGCTGCGTTCGCTGCGTGGGAGACCGGTATAGCTC 1799  
Db 1923 CTTCCTGCTGACTGACTGCGCTGCGTTCGCTGCGTGGGAGACCGGTATAGCTC 1864  
Qy 1800 ACTCAAAAGCGGTAAATACGTTATCAAGAAATCAAGGGAATACGACGAAAGAAATGT 1859  
Db 1863 ACTCAAAAGCGGTAAATACGTTATCAAGAAATCAAGGGAATACGACGAAAGAAATGT 1804  
Qy 1860 GAGCAAAAGCGCAAGAAAGCCAGAAACCGTAAAGACCGCGTTCGCTGCTTTTCC 1919  
Db 1803 GAGCAAAAGCGCAAGAAAGCCAGAAACCGTAAAGACCGCGTTCGCTGCTTTTCC 1744  
Qy 1920 ATAGCTCCGCCCCCTACAGAGCTACAAAAATCAAGCTCAAGTCAAGGTGGGAA 1979  
Db 1743 ATAGCTCCGCCCCCTACAGAGCTACAAAAATCAAGCTCAAGTCAAGGTGGGAA 1684  
Qy 1980 ACCGACAGACCTAATAAGATACAGAGGTTTCCCGTGAAGTCCCTGTGTGCGCTTC 2039  
Db 1683 ACCGACAGACCTAATAAGATACAGAGGTTTCCCGTGAAGTCCCTGTGTGCGCTTC 1624  
Qy 2040 CTGTCCGACCCTGCGCTTACCGGATACCTGTGCGCTTTCTCCCTGGGAAAGGTGG 2099  
Db 1623 CTGTCCGACCCTGCGCTTACCGGATACCTGTGCGCTTTCTCCCTGGGAAAGGTGG 1564  
Qy 2100 CGCTTTCTCATAGCTACGCTGATAGTATCTCAGTTCCGATGAGGTGCTTCGCTCAAGC 2159  
Db 1563 CGCTTTCTCATAGCTACGCTGATAGTATCTCAGTTCCGATGAGGTGCTTCGCTCAAGC 1504  
Qy 2160 TGGCTGTGTGCAAGAACCCCGCTTCAAGCCGACCGCTGCAGCTTATCCGTAATC 2219  
Db 1503 TGGCTGTGTGCAAGAACCCCGCTTCAAGCCGACCGCTGCAGCTTATCCGTAATC 1444  
Qy 2220 GTCTTAGTCAACCCCGTAAAGACAGACTTATCCGCACTGGACAGACCACTGGTAACA 2279  
Db 1443 GTCTTAGTCAACCCCGTAAAGACAGACTTATCCGCACTGGACAGACCACTGGTAACA 1384  
Qy 2280 GATTTAGCAGAGGAGATATGTAAGCGGTGCTACAGAGTTCTTGAAGTGTGGGCTAACT 2339  
Db 1383 GATTTAGCAGAGGAGATATGTAAGCGGTGCTACAGAGTTCTTGAAGTGTGGGCTAACT 1324  
Qy 2340 ACGGCTACACTAAGAGCAGATTTGTAATCTGCGCTGCTGAAGCAGTAACTCTTCG 2399  
Db 1323 ACGGCTACACTAAGAGCAGATTTGTAATCTGCGCTGCTGAAGCAGTAACTCTTCG 1264  
Qy 2400 GAAAAAGATTGTGCTTGTATCCGCAAAACAACCGCTGATGCGGTGTTT 2459

Db 1263 GAAAAAGATTGTAGCTTGTATCCGGCAAAACAACCCGCTGTAGCGGTGTTT 1204  
Qy 2460 TTGTTGCAACACACATTAATCCGCGAAGAAAAAGATCTCAAGAGATCTTTGATCT 2519  
Db 1203 TTGTTGCAACACACATTAATCCGCGAAGAAAAAGATCTCAAGAGATCTTTGATCT 1144  
Qy 2520 TTTTACGGGATCGACGCTACGTGAAAGAACTCAGTTAAGGATTTGTGTATGA 2579  
Db 1143 TTTTACGGGATCGACGCTACGTGAAAGAACTCAGTTAAGGATTTGTGTATGA 1084  
Qy 2580 GATTATAAAAAGATCTTCACTTATCTTTAAATTTAAATGAATTTAAATCAA 2639  
Db 1083 GATTATAAAAAGATCTTCACTTATCTTTAAATTTAAATGAATTTAAATCAA 1024  
Qy 2640 TCTAAGTATATATGATTAACCTGAGTGAAGTAACTCAATGCTTATCTGAGGAC 2699  
Db 1023 TCTAAGTATATATGATTAACCTGAGTGAAGTAACTCAATGCTTATCTGAGGAC 964  
Qy 2700 CTATCTACGAGATCTGTCTAATTTGTTTCAATCCATAGTTGCTGACTCCCGCTGTAGA 2759  
Db 963 CTATCTACGAGATCTGTCTAATTTGTTTCAATCCATAGTTGCTGACTCCCGCTGTAGA 904  
Qy 2760 TAACTACATACGAGAGGCTTACCATCTGGCCCAAGTGTGAAATATACCGGAGAC 2819  
Db 903 TAACTACATACGAGAGGCTTACCATCTGGCCCAAGTGTGAAATATACCGGAGAC 844  
Qy 2820 CAGGCTCACCGGCTCAGATTTATCAAGATTAACAGCCAGCCGAGAGGCGGAGC 2879  
Db 843 CAGGCTCACCGGCTCAGATTTATCAAGATTAACAGCCAGCCGAGAGGCGGAGC 784  
Qy 2880 GAAATGTCCTGCACTTTATCCGCTCCATCCAGTCTAATTAATTTGTCGGGAAGCTA 2939  
Db 783 GAAATGTCCTGCACTTTATCCGCTCCATCCAGTCTAATTAATTTGTCGGGAAGCTA 724  
Qy 2940 GAGTAAAGATTCCCGATTAATATGTTTGGCAACGTTTGGCATTTGCTACAGCATCG 2999  
Db 723 GAGTAAAGATTCCCGATTAATATGTTTGGCAACGTTTGGCATTTGCTACAGCATCG 664  
Qy 3000 TGGTGTCAAGCTGCTGTTTGTATAGCTTCAATTCAGCTCCGTTCCCAAGATCAAGC 3059  
Db 663 TGGTGTCAAGCTGCTGTTTGTATAGCTTCAATTCAGCTCCGTTCCCAAGATCAAGC 604  
Qy 3060 GAGTTAATATCCCAATTTGTGCAAAAAAGCGTTAGCTCTTCGCTCCGATCG 3119  
Db 603 GAGTTAATATCCCAATTTGTGCAAAAAAGCGTTAGCTCTTCGCTCCGATCG 544  
Qy 3120 TTGTCAAGATTAAGTGGCGAGTGTATCACTAGTGTATATGGCAGACCTGATTAAT 3179  
Db 543 TTGTCAAGATTAAGTGGCGAGTGTATCACTAGTGTATATGGCAGACCTGATTAAT 484  
Qy 3180 CTCTAATCTATGCAATCCGTAAGTGTCTTGTGATAGTGTAGTCAACCAAGT 3239  
Db 483 CTCTAATCTATGCAATCCGTAAGTGTCTTGTGATAGTGTAGTCAACCAAGT 424  
Qy 3240 CATTTGAAATAGTGTATGCGGACAGAGTGTCTTGGCCGCGCTCAATACGGGATA 3299  
Db 423 CATTTGAAATAGTGTATGCGGACAGAGTGTCTTGGCCGCGCTCAATACGGGATA 364  
Qy 3300 ATACCGCGCAGATACGAACTTTAAAGTGCATCATTTGAAAGAGTTCTTCGGGCG 3359  
Db 363 ATACCGCGCAGATACGAACTTTAAAGTGCATCATTTGAAAGAGTTCTTCGGGCG 304  
Qy 3360 GAAAACTGTCAAGATCTTACCGCTGTGAGATCCATTTGATGAATCCATCTGTGAC 3419  
Db 303 GAAAACTGTCAAGATCTTACCGCTGTGAGATCCATTTGATGAATCCATCTGTGAC 244  
Qy 3420 CCAATGATCTTACGATCTTTTACCTTCAACAGCTTTCTGGGTGAGCAAAAGAGAA 3479  
Db 243 CCAATGATCTTACGATCTTTTACCTTCAACAGCTTTCTGGGTGAGCAAAAGAGAA 184  
Qy 3480 GGCAAAAATGCGCAAAAGAGAAATAGGCGGACAGAAATGTTGAATATCTATCT 3539  
Db 183 GGCAAAAATGCGCAAAAGAGAAATAGGCGGACAGAAATGTTGAATATCTATCT 124

OY 3540 TCCTTTTCAATATATGAGCATTTATCAGGTTATTCCTCATGAGCGATATCATAT 3599  
DB 123 TCCCTTTTCAATATATGAGCATTTATCAGGTTATTCCTCATGAGCGATATCATAT 64  
OY 3600 TTGAATGATTTAGAAAATAAACAATAGGGGTTCCCGCAGCATTTCCCGAAAAGTGC 3659  
DB 63 TTGAATGATTTAGAAAATAAACAATAGGGGTTCCCGCAGCATTTCCCGAAAAGTGC 4  
OY 3660 CAC 3662  
DB 3 CAC 1  
RESULT 11  
AAQ13578  
ID AAQ13578 standard; DNA; 3681 BP.  
AC AAQ13578;  
XX 02-DEC-1991 (first entry)  
DT  
XX  
DE Plasmid pKSEB5.  
KM recombinant expression plasmid; phage phi X174; protein E;  
KW bacterial ghost; 88.  
XX  
OS Synthetic.  
XX MO9113155-A.  
PN  
XX 05-SEP-1991.  
PD  
XX 24-FEB-1990; 90DE-04005874.  
PF  
XX 24-FEB-1990; 90DE-04005874.  
PR  
XX 24-FEB-1990; 90DE-04005874.  
PA (BOEF) BOEHRINGER MANNHEIM GMBH.  
XX  
P1 Lubitz W, Szostak MP;  
DR WPI; 1991-281471/38.  
XX  
PT Carrier recombinant protein for vaccines against HIV, etc. - obtd. by  
PT expressing fusion protein gene in gram negative bacteria and gene that  
PT encodes for lytic membrane protein.  
XX  
PS Example 1; Page 33-34; 45pp; German.  
XX  
CC Plasmid pKSEB5 comprises three multiple cloning sites (mccl-3), the  
CC ampicillin resistance gene, regions of the lac operon and a partial  
CC phage phi X174 E' target sequence. This plasmid can be used as a carrier  
CC vector for the insertion of the HIV gp41 sequence of sequences encoding  
CC viral antigenic peptides, e.g. the HIV gp41 sequence was isolated as a  
CC HindIII/PvuII fragment from plasmid pHR14. The fragment also included a  
CC linker sequence and the last 45 codons of gp120. The fragment was ligated  
CC to pKSEB5 which had been cleaved with AccI and end-filled. Alternatively,  
CC a new carrier vector, denoted pMTV1 (see AAQ13576), can be constructed  
CC from pKSEB5 which has been partially digested with SmaI. It is ligated to  
CC a SmaI-DraIII fragment from pMT1 (see AAQ13577). Foreign sequences are  
CC cloned into mcs2  
XX  
SQ Sequence 3681 BP; 949 A; 917 C; 932 G; 883 T; 0 U; 0 Other;  
Query Match 72.0%; Score 2635.8; DB 2; Length 3681;  
Best Local Similarity 83.4%; Pred. No. 0;  
Matches 3073; Conservative 0; Mismatches 582; Indels 28; Gaps 6;  
OY 3 AAATGTTAGCGTTAAATTTTGTAAATTCGCGTTAAATTTTGTAAATCAGCTCAT 62  
DB 1 AAATGTTAGCGTTAAATTTTGTAAATTCGCGTTAAATTTTGTAAATCAGCTCAT 60  
OY 63 TTTTAAACAATAGCGCAATTCGCAAAATCCCTTAATCAAAAATAAGACCGAGA 122

DB 61 TTTTAAACAATAGCGCAAAATTCGCAAAATCCCTTAATCAAAAATAAGACCGAGA 120  
OY 123 TAGGGTTAGTGTGTTCCAGTTTGAACAAGATCCATTTAAAGACGTGACCTCA 182  
DB 121 TAGGGTTAGTGTGTTCCAGTTTGAACAAGATCCATTTAAAGACGTGACCTCA 180  
OY 183 ACGTCAAAAGGCGAAAACCGTCTATCAAGGCGATGCGCCCATGATCAATCACTCT 242  
DB 181 ACGTCAAAAGGCGAAAACCGTCTATCAAGGCGATGCGCCCATGATCAATCACTCT 240  
OY 243 AATCAAGTTTGTGAGGTGAGGTGCGGTAAAGCACTAAATCGAACCCTAAAGGAGCC 302  
DB 241 AATCAAGTTTGTGAGGTGAGGTGCGGTAAAGCACTAAATCGAACCCTAAAGGAGCC 300  
OY 303 CCGGATTTAGACTTGAACGGGAAAACCGCGGAACGTGCGGAGAAAGGAAAGAAAG 362  
DB 301 CCGGATTTAGACTTGAACGGGAAAACCGCGGAACGTGCGGAGAAAGGAAAGAAAG 360  
OY 363 CGAAAGAGCGGGCGCTAGAGGCGCTGCGCAAGTGAAGCGGTCAAGCTGCGCTTAACCA 422  
DB 361 CGAAAGAGCGGGCGCTAGAGGCGCTGCGCAAGTGAAGCGGTCAAGCTGCGCTTAACCA 420  
OY 423 CACCGCGCGGCTTAATGCGCGGCTACAGAGGCGGCTCCCATTTGCGCATTCAGGCTGCGCA 482  
DB 421 CACCGCGCGGCTTAATGCGCGGCTACAGAGGCGGCTCCCATTTGCGCATTCAGGCTGCGCA 480  
OY 483 ACTGTGGGAAGGCGCATCGGTGCGGGCGCTTTCGCTATTAACGCAAGCTGCGGAAGGGG 542  
DB 481 ACTGTGGGAAGGCGCATCGGTGCGGGCGCTTTCGCTATTAACGCAAGCTGCGGAAGGGG 540  
OY 543 GATGTGCTGCAAGCGCATTTAAGTGGTAACGCGAAGGTTTCCAGTCAAGAGCTTGA 602  
DB 541 GATGTGCTGCAAGCGCATTTAAGTGGTAACGCGAAGGTTTCCAGTCAAGAGCTTGA 600  
OY 603 AAACGAGCGCCAGTGAAGCGCGCTCGTTCACTCAAGTTTGAACCGGTGAGGAGCGGCG 662  
DB 601 AAACGAGCGCCAGTGAAGCGCGCTCGTTCACTCAAGTTTGAACCGGTGAGGAGCGGCG 660  
OY 663 AGACTCGCGGTGCAAAATGTGTTTAAACGCGTGAAGAGCAAGATGATGTCAGACGCG 722  
DB 661 GTGCGGCGCGCTGATGATGATGTCAGCTGCAATTCGCTGATGCGCGGATGTTA 720  
OY 723 TGCAGAAACGACGCTGATTAACCTTAAGAAATTAATCATATTGTGACGTAAAG 782  
DB 721 AGCCAGTATATACACTCGGCTATGCTGATGATGATGATGATGATGATGATGATGATG 780  
OY 783 ATATCATGCGGTAAATATGAACGATGAGATCTGTAATACGATCACTATAGGGGGAATG 842  
DB 781 GCCAACACCGGCTGACGCGGCGCTGACGAGGCTTGCTGCTCCCGCATCCCTTAACAGCA 840  
OY 843 GGTACCGGCGCCCGCTCGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 896  
DB 841 AGCTGATACGCTCCGAGAGCTGATGTCAGAGGTTTTCACCGTATCAACGAAACG 900  
OY 897 CTGACGCGCGGAGATCACTAGTTCTAAGCGGCGCCACCGCGGTGAGCTCAAGTT 956  
DB 901 CGCAGGACAGTAAAGTGGATGCTTGTGAGCAATTCGCTTAAAGTGAATGATGCTG 960  
OY 957 TTGTCCCTTAAAGAGAGGTTAATTAATCCAT-----GCGTCAATTTAGCGAGAC 1009  
DB 961 TAAAGTGTCACCTGCGGATACACCGCTTCAGTAGAGCAAGAACAAATGATGTAA 1020  
OY 1010 TATCTTTCTAGAGGTTAATCTAGCTGATCAAGATCATATCGTCCGGTCTTTTTCGCGCT 1069  
DB 1021 TTTTGAGAGAAAGATCGAGAGAAATCATATCAATTAAGAGTTGAACCTCTTTGTGTCT 1080  
OY 1070 CAGTCAATCGCCCAAGCTGCGCTATCTGCGCATCTGGGAGAGAGAACCGCTGCTTTTC 1129  
DB 1081 TCAGCATGAGGTAATCTCAATGTTGAATGCGCTTGAAGATTCGCGCAAGCTTCAATGCC 1140  
OY 1130 CCGGAGGTTGAACGCGGATGAGAAAGATTTGCCGAGATGATGCTGCTGATTTAGCT 1189



Dh 1141 TCGAGGTGACTCTAGAGATCCCGACGCTCGAGCCATTAAATATGTTTCCGTAAT 1200  
Qy 1190 TTAGGAAAAAGCAGCTTATACCATGATGATTCGGGAAAGTGTGSCATGACGCTTTAA 1249  
Db 1201 TTAGGCGCTTCCATGATGAGACAGGCCGTTT--GAAATGTGACGGGATGAACTAATAAG 1258  
Qy 1250 CGGTAACTGTTCGTTCAAGGCACCTGGGATACCAAGTTCCGCGGCTTTTCCGACACA 1309  
Db 1259 CAATGACGGCAGCAATAACTCAACAGG---AGAGAGAAAGCCAGGGATATCCACAAGT 1315  
Qy 1310 GTTCCGAGTGTCAAGCCGGAAGCCATCAGCAACCCGACATACCGGAGAGG----- 1363  
Db 1316 CCAAGCTACCATTAAGCCAAAGCCCTCAACGAGGAGACAGACAGACCGCTCATAGCAA 1375  
Qy 1364 CCGGAACCTGCGCTCGGCTGTGAGATTAAGACAGGGGTGCGGCGCTGGGATATTACGT 1423  
Db 1376 TCCAAACTTTGTTACTGTCAAGAAATCGAAATCATCTTCGGTTAAATCAAAACGGCAG 1435  
Qy 1424 CAGCGAGAGCGGGTAT---CTTGGCTGATGCCGAGAAATGACATGGAATACCCCGTG 1479  
Db 1436 AAGCTGAATGAGAAATCGACTCGAGGGGGGGCCGGTACCAGCTTTTGTTCCTTTA 1495  
Qy 1480 AGTTACCCGGGCGGCGCTTGGCGTAAATCATGCTGATAGCTGTTTCTGTGTGAATTG 1539  
Db 1496 GTGAGGTTAATTCAGCTTGCGCTGTAATCATGCTATGCTGTTCTGTGTGAATTG 1555  
Qy 1540 TTATCCGCTCAAAATTCACACAAACATACAGCCGGAGCATTAAGTGTAAAGCTGGGG 1599  
Db 1556 TTATCCGCTCAAAATTCACACAAACATAGAGCCGGAGCATTAAGTGTAAAGCTGGGG 1615  
Qy 1600 TGCCTAATGATGAGCTAATCTCAATTAATTCGCTTGGCTGCTCACTGCCGCTTTCAGTC 1659  
Db 1616 TGCCTAATGATGAGCTAATCTCAATTAATTCGCTTGGCTGCTCACTGCCGCTTTCAGTC 1675  
Qy 1660 GGGAAAAGCTGTGTCCAGCTGCAATTAATGAAATCGGACCGGGGAGAGGCGGTTT 1719  
Db 1676 GGGAAAAGCTGTGTCCAGCTGCAATTAATGAAATCGGACCGGGGAGAGGCGGTTT 1735  
Qy 1720 GGGTATGGGCGCTTTCGCTTCCGCTTCTGCTCACTGACTGTGGCTCGTGTGCTGGCT 1779  
Db 1736 GGGTATGGGCGCTTTCGCTTCCGCTTCTGCTCACTGACTGTGGCTCGTGTGCTGGCT 1795  
Qy 1780 GGGGAGACGGTATCAGCTCACTCAAAAGCGGTAATCGGTTATCCACAAATCAGGGGA 1839  
Db 1796 GGGGAGACGGTATCAGCTCACTCAAAAGCGGTAATCGGTTATCCACAAATCAGGGGA 1855  
Qy 1840 TAAAGCAGGAGAAACATGTAGCAAAAGCCAGCAAAAGCCAGGAAACGTTAAAGGC 1899  
Db 1856 TAAAGCAGGAGAAACATGTAGCAAAAGCCAGCAAAAGCCAGGAAACGTTAAAGGC 1915  
Qy 1900 CGGCTTGTGCGCTTTTTCATAGGCTCCGCCCCCGTGAAGACATCAAAATGAGC 1959  
Db 1916 CGGCTTGTGCGCTTTTTCATAGGCTCCGCCCCCGTGAAGACATCAAAATGAGC 1975  
Qy 1960 CTAAAGTCAAGGTGGGAAACCCGACAGGACTATAAAGATACAGGCTTTCCTCCCTG 2019  
Db 1976 CTAAAGTCAAGGTGGGAAACCCGACAGGACTATAAAGATACAGGCTTTCCTCCCTG 2035  
Qy 2020 AAGCTCCCTGTGCGCTCTCTGTTCGACCCGCGGCTTAACGGATACCTGTCCGCTT 2079  
Db 2036 AAGCTCCCTGTGCGCTCTCTGTTCGACCCGCGGCTTAACGGATACCTGTCCGCTT 2095  
Qy 2080 TCTCCCTTGGGAAAGCGTGGCGCTTTCATAGCTCAAGCTGTAGGTATCTCAAGTTGGT 2139  
Db 2096 TCTCCCTTGGGAAAGCGTGGCGCTTTCATAGCTCAAGCTGTAGGTATCTCAAGTTGGT 2155  
Qy 2140 GTAGTGTGTGCTCAAGCTGGGCTGTGTGACAAACCCCGCTTACGCCGACCGCTG 2199  
Db 2156 GTAGTGTGTGCTCAAGCTGGGCTGTGTGACAAACCCCGCTTACGCCGACCGCTG 2215  
Qy 2200 CGCTTATCCGGTAATCTGTCTTGTAGTCAACCCGGTAAAGACAGCATTAATGCCACT 2259  
Db 2216 CGCTTATCCGGTAATCTGTCTTGTAGTCAACCCGGTAAAGACAGCATTAATGCCACT 2275

---

Qy 2260 GGCAGACCACTGTGTAACAGATTAGCAGAGCGGTATGTAGCGGCTGTACAGATT 2319  
Db 2276 GGCAGACCACTGTGTAACAGATTAGCAGAGCGGTATGTAGCGGCTGTACAGATT 2335  
Qy 2320 CTGTAAGTGTGGCTTAACCTACAGCTACACTAGAAAGACAGTATTTGTATCTGCGCTC 2379  
Db 2336 CTGTAAGTGTGGCTTAACCTACAGCTACACTAGAAAGACAGTATTTGTATCTGCGCTC 2395  
Qy 2380 GCTAAGCCAGTTAATCTTGGGAAAAAGTGTGTACTCTTGTATCCGGCAACAAACAC 2439  
Db 2396 GCTAAGCCAGTTAATCTTGGGAAAAAGTGTGTACTCTTGTATCCGGCAACAAACAC 2455  
Qy 2440 CGCTGTAGTGGGT 2499  
Db 2456 CGCTGTAGTGGGT 2515  
Qy 2500 TCAAGAAATCTTTGATCTTTCTAAGGGGTCTGACGCTCAGTGAACGAAACCTCACG 2559  
Db 2516 TCAAGAAATCTTTGATCTTTCTAAGGGGTCTGACGCTCAGTGAACGAAACCTCACG 2575  
Qy 2560 TTAAGGATTTTGTATGATAGATTAACAAAAAGATCTTCACTAGATCTTTTAAATTA 2619  
Db 2576 TTAAGGATTTTGTATGATAGATTAACAAAAAGATCTTCACTAGATCTTTTAAATTA 2635  
Qy 2620 AAAATGAATTTTAAATCAATCTAAGTATATATAGTAAACCTTGTCTGACATTAACA 2679  
Db 2636 AAAATGAATTTTAAATCAATCTAAGTATATATAGTAAACCTTGTCTGACATTAACA 2695  
Qy 2680 ATGCTTATCAGTAGGACCTAATCTAGCATCTGTATTTGCTTATCATCAATAGTTGC 2739  
Db 2696 ATGCTTATCAGTAGGACCTAATCTAGCATCTGTATTTGCTTATCATCAATAGTTGC 2755  
Qy 2740 CTGACTCCCGCGGTGATTAATCAAGTACAGTACGGAAGGCTTACATCTGAGCCAGTGC 2799  
Db 2756 CTGACTCCCGCGGTGATTAATCAAGTACAGTACGGAAGGCTTACATCTGAGCCAGTGC 2815  
Qy 2800 TGCATGTATCCGAGAACCCACGCTCAGGCTCAGATTATTCAGCAATTAACCAAGCC 2859  
Db 2816 TGCATGTATCCGAGAACCCACGCTCAGGCTCAGATTATTCAGCAATTAACCAAGCC 2875  
Qy 2860 AGCCGAAAGGGCGAGGCGAAGTGTCTCTGCACTTTATCCGCTCATCAAGTCTAT 2919  
Db 2876 AGCCGAAAGGGCGAGGCGAAGTGTCTCTGCACTTTATCCGCTCATCAAGTCTAT 2935  
Qy 2920 TAAATTTGCCGGGAACCTAGATTAAGTGTGCGCAATTAATAGTTTCCGCAAGTGT 2979  
Db 2936 TAAATTTGCCGGGAACCTAGATTAAGTGTGCGCAATTAATAGTTTCCGCAAGTGT 2995  
Qy 2980 TGCAATGTACAGGCAATCGTGTGTACGCTGTGCTGTGTGTGTGTGTGTGTGTGTGTGT 3039  
Db 2996 TGCAATGTACAGGCAATCGTGTGTACGCTGTGCTGTGTGTGTGTGTGTGTGTGTGTGT 3055  
Qy 3040 CGGTTCCAACGATCAAGCGAGTATCATGATCCCCCAATGTTGTGCAAAAAAGCGTTAG 3099  
Db 3056 CGGTTCCAACGATCAAGCGAGTATCATGATCCCCCAATGTTGTGCAAAAAAGCGTTAG 3115  
Qy 3100 CTGCTTGGGTCCCGAATGCTTGTCAAGATTAAGTTGGCCGAGTGTATCATCATAGT 3159  
Db 3116 CTGCTTGGGTCCCGAATGCTTGTCAAGATTAAGTTGGCCGAGTGTATCATCATAGT 3175  
Qy 3160 TATGAGACATGCAATATCTTACTGTATGCAATGCAATCCGTTAGATGCTTTTGTGAC 3219  
Db 3176 TATGAGACATGCAATATCTTACTGTATGCAATGCAATCCGTTAGATGCTTTTGTGAC 3235  
Qy 3220 TGTGATGATCAACCAAGTATCTGAAATAGTATGCGGACAGAGTTGCTCTTG 3279  
Db 3236 TGTGATGATCAACCAAGTATCTGAAATAGTATGCGGACAGAGTTGCTCTTG 3295  
Qy 3280 CCCGCGTCAATACGGGATTAATACCGGCAATAGAGAACTTTAAAGTGTCTATCAT 3339  
Db 3296 CCCGCGTCAATACGGGATTAATACCGGCAATAGAGAACTTTAAAGTGTCTATCAT 3355



Oy 3340 TGAAGAACTCTTGGGGGCGAAAACTCTCAAGACTCTTACCGCTGTGAGTCAGTTC 3399  
 Db 3356 TGAAGAACTCTTGGGGGCGAAAACTCTCAAGACTCTTACCGCTGTGAGTCAGTTC 3415  
 Oy 3400 GATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTACTTCACACAGCTTC 3459  
 Db 3416 GATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTACTTCACACAGCTTC 3475  
 Oy 3460 TGGGTGAGCAAAAACAGAGAGCAAAATGCGCAAAAAAGGAAATAGCGCGACAGAA 3519  
 Db 3476 TGGGTGAGCAAAAACAGAGAGCAAAATGCGCAAAAAAGGAAATAGCGCGACAGAA 3535  
 Oy 3520 ATGTGATACCTACTCTCTCTTTCATATATTATGAGCATTTATCAGGGTTATG 3579  
 Db 3536 ATGTGATACCTACTCTCTCTTTCATATATTATGAGCATTTATCAGGGTTATG 3595  
 Oy 3580 TCTCATGAGCGGATACATATTTGATGATGATTTAGAAAAATAAACAATAGGGTTCCGCG 3639  
 Db 3596 TCTCATGAGCGGATACATATTTGATGATGATTTAGAAAAATAAACAATAGGGTTCCGCG 3655  
 Oy 3640 CACATTTTCCCGAAAAAGTCCAC 3662  
 Db 3656 CACATTTTCCCGAAAAAGTCCAC 3678

RESULT 12

AAQ13576 ID AAQ13576 standard; DNA; 5314 BP.

AC AAQ13576;

DT 02-DEC-1991 (first entry)

DE Plasmid pMTV1 containing 3 multiple cloning sites.

KM recombinant expression plasmid; phage phi X174; protein E;

XX bacterial ghost; ss.

OS Synthetic.

PN MO9113155-A.

PD 05-SEP-1991.

PF 24-FEB-1990; 90DE-04005874.

PR 24-FEB-1990; 90DE-04005874.

PA (BOEF) BOEHRINGER MANNHEIM GMBH.

PI Lubitz W, Szostak MP;

DR WPI; 1991-281471/38.

PT Carrier recombinant protein for vaccines against HIV, etc. - obt'd. by

PT expressing fusion protein gene in gram negative bacteria and gene that

PS encodes for lytic membrane protein.

XX Example 10; Page 28-29; 45pp; German.

CC The carrier plasmid pMTV1 comprises the phage phi X174 E protein gene,

CC the ampicillin resistance gene, the lambda c1857 repressor gene and P(R)

CC from the lambda promoter/operator system. There are also lac operon

CC sequences and three multiple cloning sites. A foreign gene can be

CC inserted into the second multiple cloning site (mcs2). The protein

CC encoded by the foreign gene is especially a viral antigenic peptide, e.g.

CC of HIV, HBV or EBV. See also AAQ13577 and AAQ13578

CC XX Sequence 5314 BP; 1438 A; 1276 C; 1310 G; 1290 T; 0 U; 0 Other;

CC Query Match 71.6%; Score 2622; DB 2; Length 5314;

CC Best Local Similarity 83.4%; Pred. No. 0;

CC Matches 3058; Conservative 0; Mismatches 580; Indels 28; Gaps 6;

Oy 20 ATTTGTTAAATTCGGTTAAATTTTGTAAATCAGCTATTTTAAACCAATAGGCC 79  
 Db 1651 ATTTGTTAAATTCGGTTAAATTTTGTAAATCAGCTATTTTAAACCAATAGGCC 1710  
 Oy 80 GAAATCGGCAAAATCCCTTATTAATCAAAAGATAGACCGATAGGGTTAGTGTG 139  
 Db 1711 GAAATCGGCAAAATCCCTTATTAATCAAAAGATAGACCGATAGGGTTAGTGTG 1770  
 Oy 140 CCAAGTTGAAACAAAGTCCATTAATTAAGAACGTGACCTTCAACCTCAAGGGCGAAAA 199  
 Db 1771 CCAAGTTGAAACAAAGTCCATTAATTAAGAACGTGACCTTCAACCTCAAGGGCGAAAA 1830  
 Oy 200 ACCGTTATCAGGGCGATGCGCCCTACGTAAACATCAACCTTATCAAGTTTGGGG 259  
 Db 1831 ACCGTTATCAGGGCGATGCGCCCTACGTAAACATCAACCTTATCAAGTTTGGGG 1890  
 Oy 260 TCGAGGTGCGCTAAGACCTAAATCGGAACCTTAAAGGAGCCCGATTTAGAGCTTGA 319  
 Db 1891 TCGAGGTGCGCTAAGACCTAAATCGGAACCTTAAAGGAGCCCGATTTAGAGCTTGA 1950  
 Oy 320 CGGGGAAAGCCGCGAAAGTGGCGAGAAAGGAAAGGAAAGGAAAGGAGCGCGCT 379  
 Db 1951 CGGGGAAAGCCGCGAAAGTGGCGAGAAAGGAAAGGAAAGGAAAGGAGCGCGCT 2010  
 Oy 380 AGGCGCTGGCAATGTAGCGGTACCGCTGCGCTAACACACACACCCCGCGCTTAAT 439  
 Db 2011 AGGCGCTGGCAATGTAGCGGTACCGCTGCGCTAACACACACACCCCGCGCTTAAT 2070  
 Oy 440 GCGCGCTGACAGGGCGGGTCCCATTTGCGCAATTCAGGCGGCACTGTGGGAAAGGGGA 499  
 Db 2071 GCGCGCTGACAGGGCGGGTCCCATTTGCGCAATTCAGGCGGCACTGTGGGAAAGGGGA 2130  
 Oy 500 TCGGTGCGGGCTCTTGTGCTAATTAAGCGAGCTGCGGAAAGGGGATGTGTCAAGCGGA 559  
 Db 2131 TCGGTGCGGGCTCTTGTGCTAATTAAGCGAGCTGCGGAAAGGGGATGTGTCAAGCGGA 2190  
 Oy 560 TTAAGTTGGGTACCGCAAGGTTTCCAGTACGACGTTGTAAACGACCGCACTGAG 619  
 Db 2191 TTAAGTTGGGTACCGCAAGGTTTCCAGTACGACGTTGTAAACGACCGCACTGAG 2250  
 Oy 620 CGCGCTGCTTCAATTCAGTTTGAACCGGTGAGAGACCGGCGAGACTCGCGGTGCAAT 679  
 Db 2251 TTAAGTTGGGTACCGCAAGGTTTCCAGTACGACGTTGTAAACGACCGCACTGAG 2310  
 Oy 680 GTGTTTAAACCGGTGAGAGCAAGATGCTGCACACGCTGCAGAACACCGCACTA 739  
 Db 2311 ATGTTGCACTCTCAGTACATCTGCTGATGCCGCAATGTTAAGCCAGTATACACTC 2370  
 Oy 740 GATTAACTTGAAGAAATATCATATTGTGACGTAGCTTAAAGATATCATGCTTAAAT 799  
 Db 2371 GATTAACTTGAAGAAATATCATATTGTGACGTAGCTTAAAGATATCATGCTTAAAT 800  
 Oy 800 TGACGCATGGATGTGTAATACGACTACATAGGGGAAATGGGTACCGGGCCCCCT 859  
 Db 2431 TGACGCATGGATGTGTAATACGACTACATAGGGGAAATGGGTACCGGGCCCCCT 859  
 Oy 2431 CGCCCTAGCGGCTTGTGCTGCTCCGCGATCCGCTTACAGCAAGCTGTGACCTTCCG 2490  
 Db 860 CGAGTCGACGATACGACTTAAAGTTTCCGACGCGCGGGGATC 913  
 Oy 2491 CGAGTCGACGATACGACTTAAAGTTTCCGACGCGCGGGGATC 913  
 Db 914 CACTAGTTTGAAGCGGCGCGACCGGCTGAGAGCTTCAAGCTTAAAGGAGTGC 973  
 Oy 2551 GGAATGCTTGTGAGCAATTCGCTCCCTTAAGTGAATTCGTTAAGTGTGCTGTC 2610  
 Db 974 GGAATGCTTGTGAGCAATTCGCTCCCTTAAGTGAATTCGTTAAGTGTGCTGTC 2610  
 Oy 2611 GGAATGCTTGTGAGCAATTCGCTCCCTTAAGTGAATTCGTTAAGTGTGCTGTC 2670  
 Db 1027 TCTAGTCGATCAGAGATCATATGCTCGGGTCTTTTTCGAGCTCAGTACGCCAAGCT 1086  
 Oy 2671 CGAGGAAATCATATCAATTAAGAGTTGAACTCTTTGTTGCTTTCGACATGGGTAATCT 2730

OY 1087 GGGCTATCTGGGCATCGGAGAGAAAGCCCGTCTTTCCCGCAGGTTGAAGCGG 1146  
| | | | |  
Db 2731 CATGTTGAATGGCCCTTAAGATCCGGCAAGCTTGCATGCTGCAAGTGAATCTAGA 2790  
| | | | |  
OY 1147 CATGAAAGATTGTCGAGATGA CTGCTGCTGATGACGTTGAGGAAAGCGACGT 1206  
| | | | |  
Db 2791 GGAATCCCAACGCTCGACGCATTAATATGTTTCCGTAATTCAGCGCTTCATGAT 2850  
| | | | |  
OY 1207 TTACCAATGATGATGGGAAGGTGTGGCATGCAAGCCTTTAAAGTGAATGTTGCTTC 1266  
| | | | |  
Db 2851 GAGACAGCGCTTTT--GAATGTTGACGGATGAACAATAATGACATGACGACGAAT 2908  
| | | | |  
OY 1267 AGGCCACTGGGATACAGATTCTGCGCGCTTTCCGACACAGTTCCGATGCTAGCC 1326  
| | | | |  
Db 2909 AACTCAACAGG---AGCAGAAAGCAGAGGATATCCCAAAAGTCCAGCCATCAATAACG 2965  
| | | | |  
OY 1327 CGAAGCGATGAGCAACCCGAAACAATCCGCGCAG-----CCGAAATCGCGTGGCG 1380  
| | | | |  
Db 2966 CAAAGCTTACAGACGACGACGAGCAGAGCGGTCAGTAGCAATCCAACTTTGTTACTC 3025  
| | | | |  
OY 1381 GTTGCAGATTAAATGACAGCGGTGCGCGCTGGGATTAATTCGTACAGCAGAGCGGAT- 1439  
| | | | |  
Db 3026 GTCAAGAAATCGAAATCATCTTCGTTAAATCCAAAGCGCAGAAAGCTGAATGAGAAAT 3085  
| | | | |  
OY 1440 ---CCTGGCTGAGTCCGCGAAGAAATGGAATGATACCCCGTGAATTACCGCGGGGCG 1496  
| | | | |  
Db 3086 CGACTCGAGGGGGGCGCGTACCGAGCTTTTGTCTCTTAATGAGGTTAATTTCCGA 3145  
| | | | |  
OY 1497 GCTTGGCGTAATCATGCTCATAGCTGTTTCTGTGTGAAATTTGTTATCCGCTCACAATTC 1556  
| | | | |  
Db 3146 GCTTGGCGTAATCATGCTCATAGCTGTTTCTGTGTGAAATTTGTTATCCGCTCACAATTC 3205  
| | | | |  
OY 1557 CACAAACAATAGAGCGCGAAGCAATTAAGTTAAAGCTGGGGTGCCTAATGATGAGCT 1616  
| | | | |  
Db 3206 CACAAACAATAGAGCGCGAAGCAATTAAGTTAAAGCTGGGGTGCCTAATGATGAGCT 3265  
| | | | |  
OY 1617 AACTCAATTAATGCGTTGCGCTCACTGCGCGCTTCCAGTCGGGAAACCTGTGCTGCC 1676  
| | | | |  
Db 3266 AACTCAATTAATGCGTTGCGCTCACTGCGCGCTTCCAGTCGGGAAACCTGTGCTGCC 3325  
| | | | |  
OY 1677 AGCTGATTAATGAAATCCGCCAAGCGCGGGAGAGCGGTTTCCGTAATGGCGCTTT 1736  
| | | | |  
Db 3326 AGCTGATTAATGAAATCCGCCAAGCGCGGGAGAGCGGTTTCCGTAATGGCGCTTT 3385  
| | | | |  
OY 1737 CGGCTTCTCGCTCACTGCTGCTGCGCTCGGTGCTTCCGCGCGCAGCGGTAACG 1796  
| | | | |  
Db 3386 CGGCTTCTCGCTCACTGCTGCTGCGCTCGGTGCTTCCGCGCGCAGCGGTAACG 3445  
| | | | |  
OY 1797 CTCACTCAAAAGCGGTAATAGGTTATCCACAGATCAGGGGATTAACGAGAAAGACA 1856  
| | | | |  
Db 3446 CTCACTCAAAAGCGGTAATAGGTTATCCACAGATCAGGGGATTAACGAGAAAGACA 3505  
| | | | |  
OY 1857 TGTGAGCAAAAGCGCAGCAAAAGCGCAGAAACGTTAAAGCGCGGTTGCTGGCGTTT 1916  
| | | | |  
Db 3506 TGTGAGCAAAAGCGCAGCAAAAGCGCAGAAACGTTAAAGCGCGGTTGCTGGCGTTT 3565  
| | | | |  
OY 1917 TCCATAGAGCTCCGCCCCCTGACAGCATCAAAAATGACAGCTCAAGTCAGAGTGGC 1976  
| | | | |  
Db 3566 TCCATAGAGCTCCGCCCCCTGACAGCATCAAAAATGACAGCTCAAGTCAGAGTGGC 3625  
| | | | |  
OY 1977 GAAACCGGACAGGACTATAAGATACAGAGCGTTTCCCGTGGAGCTCCCTCGTGCCT 2036  
| | | | |  
Db 3626 GAAACCGGACAGGACTATAAGATACAGAGCGTTTCCCGTGGAGCTCCCTCGTGCCT 2096  
| | | | |  
OY 2037 CTCCTGTTCCGACCCCGCGCTTACCGGATACCTGTCCGCTTTCTCCCTCGGAAAGCG 2096  
| | | | |  
Db 3686 CTCCTGTTCCGACCCCGCGCTTACCGGATACCTGTCCGCTTTCTCCCTCGGAAAGCG 3745  
| | | | |  
OY 2097 TGGCGCTTCTCATAGCTCAAGCTGTAGGTAATCTCAGTTGGTGTAGGTGTTGCTCCA 2156  
| | | | |  
Db 3746 TGGCGCTTCTCATAGCTCAAGCTGTAGGTAATCTCAGTTGGTGTAGGTGTTGCTCCA 3805  
| | | | |  
OY 2157 AGCTGGGCTGTGTGACGAACCCCGCTTACCGCGAGCCGCTGCGCTTAATCGGTAACCT 2216  
| | | | |

Db 3806 AGCTGGGCTGTGTGACGAACCCCGCTTACGCCCGACCGCTCGCTTAATCCGTAACCT 3865  
| | | | |  
OY 2217 ATGCTTGAATCCAAACCCGTAAGACACGACTTAATGCACTGAGGAGGACACTGTA 2276  
| | | | |  
Db 3866 ATGCTTGAATCCAAACCCGTAAGACACGACTTAATGCACTGAGGAGGACACTGTA 3925  
| | | | |  
OY 2277 ACAGATTTAGACAGGAGTATGTAAGCGGCTTCAAGATTCTTGAAGTGTGCTTA 2336  
| | | | |  
Db 3926 ACAGATTTAGACAGGAGTATGTAAGCGGCTTCAAGATTCTTGAAGTGTGCTTA 3985  
| | | | |  
OY 2337 ACTAGCGCTACCTAGAGGACAGTAATTTGTATCTGCGCTCTGCGAAGCGATTAACCT 2396  
| | | | |  
Db 3986 ACTAGCGCTACCTAGAGGACAGTAATTTGTATCTGCGCTCTGCGAAGCGATTAACCT 4045  
| | | | |  
OY 2397 TCGGAAAAAGATTGGTGTAGCTCTTAATCCGGGAAACAAACCAACGCTGTAGCGGTGTT 2456  
| | | | |  
Db 4046 TCGGAAAAAGATTGGTGTAGCTCTTAATCCGGGAAACAAACCAACGCTGTAGCGGTGTT 4105  
| | | | |  
OY 2457 TTTTGTGTTGCAAGCAGATTTAGCGCAGAAAAAGATCTCAAGAAAGATCTTTGA 2516  
| | | | |  
Db 4106 TTTTGTGTTGCAAGCAGATTTAGCGCAGAAAAAGATCTCAAGAAAGATCTTTGA 4165  
| | | | |  
OY 2517 TCTTTTCTACGGGCTGTAGCGCTGAGGAAAGAAACCTCAAGATTGTTGCTCA 2576  
| | | | |  
Db 4166 TCTTTTCTACGGGCTGTAGCGCTGAGGAAAGAAACCTCAAGATTGTTGCTCA 4225  
| | | | |  
OY 2577 TGAATTAATCAAAAAGATCTTCACTAGATCTTTTAAATTAATAATGAAATTTAAAT 2636  
| | | | |  
Db 4226 TGAATTAATCAAAAAGATCTTCACTAGATCTTTTAAATTAATAATGAAATTTAAAT 4285  
| | | | |  
OY 2637 CAATCAAAATTAATTAAGTAAGTGTGTGACAGTTCAATGCTTAATCAATGAGG 2696  
| | | | |  
Db 4286 CAATCAAAATTAATTAAGTAAGTGTGTGACAGTTCAATGCTTAATCAATGAGG 4345  
| | | | |  
OY 2697 CACCTAATCAGCGATCTGTCTAATTTGTTCAATCAATGTTGCTGACCTCCCGTGT 2756  
| | | | |  
Db 4346 CACCTAATCAGCGATCTGTCTAATTTGTTCAATCAATGTTGCTGACCTCCCGTGT 4405  
| | | | |  
OY 2757 AGATTAATCAAGTAAGGAGGCTTAACATTTGGGCCCCAGTGTGATGATACCGGAG 2816  
| | | | |  
Db 4406 AGATTAATCAAGTAAGGAGGCTTAACATTTGGGCCCCAGTGTGATGATACCGGAG 4465  
| | | | |  
OY 2817 ACCGAGCTCAACGGCTCCAGATTTATCAGCAATTAACAGCAGACCGGAAAGCGCGAGC 2876  
| | | | |  
Db 4466 ACCGAGCTCAACGGCTCCAGATTTATCAGCAATTAACAGCAGACCGGAAAGCGCGAGC 4525  
| | | | |  
OY 2877 GCAGAAATGCTCTGCACTTAATCCGCTCCATCTAATTAATTTGTCGGGAG 2936  
| | | | |  
Db 4526 GCAGAAATGCTCTGCACTTAATCCGCTCCATCTAATTAATTTGTCGGGAG 4585  
| | | | |  
OY 2937 CTGAGATTAAGTGTGCGCAATTAATGTTGCGCAAGTGTGGCATTTGCTACAGGCA 2996  
| | | | |  
Db 4586 CTGAGATTAAGTGTGCGCAATTAATGTTGCGCAAGTGTGGCATTTGCTACAGGCA 4645  
| | | | |  
OY 2997 TCGTGTGTGACGCTCGTGTGTTGGTAAGCTTCAATCAAGTCCGTTCCAAAGATCA 3056  
| | | | |  
Db 4646 TCGTGTGTGACGCTCGTGTGTTGGTAAGCTTCAATCAAGTCCGTTCCAAAGATCA 4705  
| | | | |  
OY 3057 GCGAGTTTACATGATCCCCCATGTTGCAAAAAGCGGTTAGCTCTTGGGCTCCGCA 4765  
| | | | |  
Db 4706 GCGAGTTTACATGATCCCCCATGTTGCAAAAAGCGGTTAGCTCTTGGGCTCCGCA 4825  
| | | | |  
OY 3117 TCGTGTGAGAAATGTTGGCCGCAAGTATCACTCATGTTATGAGCAGCAGTGCATA 3176  
| | | | |  
Db 4766 TCGTGTGAGAAATGTTGGCCGCAAGTATCACTCATGTTATGAGCAGCAGTGCATA 4825  
| | | | |  
OY 3177 ATTCTCTTAATCTGTATGCAATCCGTATAGATGCTTTCTGTGACTGGTGAATCTCAACA 3236  
| | | | |  
Db 4826 ATTCTCTTAATCTGTATGCAATCCGTATAGATGCTTTCTGTGACTGGTGAATCTCAACA 4885  
| | | | |  
OY 3237 AGTCATCTGAGAAATGTTATGCGGCAACGATTTGCTTGGCGGCTCAATACGGG 3296  
| | | | |

Db 4886 AGCATCTCTGAGTAATGATGATGAGGCGACCGAGTGTCTCTTGCCCGGCTCAATAGCGG 4945  
 Qy 3297 ATATATCCGCGCCACATGACGAACTTTAAAGTCTCATTTGGAAGAGCTTCTTGG 3356  
 Db 4946 ATATATCCGCGCCACATGACGAACTTTAAAGTCTCATTTGGAAGAGCTTCTTGG 5005  
 Qy 3357 GCGGAAACTCTCAAGGATCTTACCGGCTTGAGATGAGTCAATGTAACCACTGCTG 3416  
 Db 5006 GCGGAAACTCTCAAGGATCTTACCGGCTTGAGATGAGTCAATGTAACCACTGCTG 5065  
 Qy 3417 CACCAACTGATCTTCAAGCATCTTTTACCTTACCAAGGCTTCTGAGTGAAGCAAAACAG 3476  
 Db 5066 CACCAACTGATCTTCAAGCATCTTTTACCTTACCAAGGCTTCTGAGTGAAGCAAAACAG 5125  
 Qy 3477 GAAGGCAAAATGCGCGCAAAAGGAAATAGGCGACACGGAATGTTGAATCTCATAC 3536  
 Db 5126 GAAGGCAAAATGCGCGCAAAAGGAAATAGGCGACACGGAATGTTGAATCTCATAC 5185  
 Qy 3537 TCTTCTTTTCAATATATATGAGCATTTATCAGGGTATTTGTCTCATGAGCGGATCA 3596  
 Db 5186 TCTTCTTTTCAATATATATGAGCATTTATCAGGGTATTTGTCTCATGAGCGGATCA 5245  
 Qy 3597 TATTTGATGATTTAGAAATTAACAATAGGGGTTCCGCGACATTTCCCGAAAG 3656  
 Db 5246 TATTTGATGATTTAGAAATTAACAATAGGGGTTCCGCGACATTTCCCGAAAG 5305  
 Qy 3657 TGCCAC 3662  
 Db 5306 TGCCAC 5311

RESULT 13  
 AAV14340  
 ID AAV14340 standard; DNA; 3699 BP.  
 XX AAV14340;  
 AC  
 XX  
 DT 19-MAY-1998 (first entry)  
 DE Plasmid pBSGFP expressing green fluorescent protein.  
 XX  
 KM Green fluorescent protein; GFP; blue fluorescent protein; BFP; jellyfish;  
 KM genetic engineering marker; gene therapy; plasmid; ss.  
 OS Synthetic.  
 XX  
 PN W09742320-A1.  
 PD 13-NOV-1997.  
 XX  
 PF 07-MAY-1997; 97WO-US007625.  
 XX  
 PR 08-MAY-1996; 96US-00646538.  
 XX  
 PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
 XX  
 PI Pavlakis GN, Gaitanaris GA, Stauber RH, Vournakis JN;  
 XX  
 DR WPI; 1997-558982/51.  
 XX  
 PT New nucleic acid encoding proteins of Aequorea victoria with increased  
 PT fluorescence - useful as markers for detecting cellular transformation,  
 PT subcellular localization of proteins, for assessing gene therapy,  
 PT mutagenicity etc.  
 XX  
 PS Example; Page 67-68; 105pp; English.  
 XX  
 CC This sequence is a plasmid used to express the mutated green fluorescent  
 CC protein (GFP) of Aequorea victoria coding sequence of the invention. The  
 CC DNA of the invention that encodes a protein that is (a) a protein that  
 CC has been at position 65 (and optionally Thr at 168 and optionally further  
 CC Cys at 66) and has cellular fluorescence at least 5 times that of wild-  
 CC type GFP or (b) is a blue fluorescent protein (BFP) with His at position

CC 67 and also at least one of Leu at 65 and Ala at 164, and has cellular  
 CC fluorescence at least 5 times that of BFP (Tyr67 to His). The nucleic  
 CC acids can be used as markers in genetic engineering and gene therapy.  
 CC They may also be used to detect and characterize regulatory and coding  
 CC sequence elements that control subcellular expression and targeting of  
 CC proteins. Typical applications are monitoring targeting and transport of  
 CC proteins in cells; assessment of gene therapy procedures; in diagnosis  
 CC (when expressed under control of a promoter induced by a particular  
 CC analyte); assessment of mutagenicity of compounds; and for drug screening  
 CC (where expression is controlled by the promoter of a target gene),  
 CC particularly for antiviral or antiparasitic agents  
 XX  
 SO Sequence 3699 BP; 942 A; 884 C; 898 G; 975 T; 0 U; 0 Other;

Query Match 71.6%; Score 2620.6; DB 2; Length 3699;  
 Best Local Similarity 83.3%; Pred. No. 0;  
 Matches 3078; Conservative 0; Mismatches 579; Indels 40; Gaps 7;

Qy 3 AAATTGTAAGGCTTAATATTTTGTAAATTTGCGTTAAATTTTGTAAATTCAGCTCAT 62  
 Db 3 AAATTGTAAGGCTTAATATTTTGTAAATTTGCGTTAAATTTTGTAAATTCAGCTCAT 62  
 Qy 63 TTTTAAACAATAGCGCGAAATTCGCAAAATCCCTTAAATCAAAAGATAGACCGAGA 122  
 Db 63 TTTTAAACAATAGCGCGAAATTCGCAAAATCCCTTAAATCAAAAGATAGACCGAGA 122  
 Qy 123 TAGGGTGAAGTGTGTTCCAGTTTGAACAAGAGTCCATTTAAAGAACTGGAATCTCA 182  
 Db 123 TAGGGTGAAGTGTGTTCCAGTTTGAACAAGAGTCCATTTAAAGAACTGGAATCTCA 182  
 Qy 183 ACCGTAAGGCGGCAAAACCGTCTATCAAGGCGGATGCGCACTGACCAATCAACCT 242  
 Db 183 ACCGTAAGGCGGCAAAACCGTCTATCAAGGCGGATGCGCACTGACCAATCAACCT 242  
 Qy 243 AATCAAGTTTGTGGGTCGAGGTCGCGTAAGCACTAAATCGGAACCTTAAAGGAGCC 302  
 Db 243 AATCAAGTTTGTGGGTCGAGGTCGCGTAAGCACTAAATCGGAACCTTAAAGGAGCC 302  
 Qy 243 AATCAAGTTTGTGGGTCGAGGTCGCGTAAGCACTAAATCGGAACCTTAAAGGAGCC 302  
 Db 243 AATCAAGTTTGTGGGTCGAGGTCGCGTAAGCACTAAATCGGAACCTTAAAGGAGCC 302  
 Qy 303 CCCGATTTAGAGCTTGAACGCGGAAAGCGCGGAAAGCTGCGAGAAAGGAAAG 362  
 Db 303 CCCGATTTAGAGCTTGAACGCGGAAAGCGCGGAAAGCTGCGAGAAAGGAAAG 362  
 Qy 362 CGAAAGAGGCGGCGCTAGAGGCGGCTGAGGCGGCTGAGGCGGCTGAGGCGGCTG 422  
 Db 362 CGAAAGAGGCGGCGCTAGAGGCGGCTGAGGCGGCTGAGGCGGCTGAGGCGGCTG 422  
 Qy 423 CACCGCGCGGCTTAATAGCGCGGCTGAGGCGGCTGAGGCGGCTGAGGCGGCTG 479  
 Db 423 CACCGCGCGGCTTAATAGCGCGGCTGAGGCGGCTGAGGCGGCTGAGGCGGCTG 479  
 Qy 480 GCAACTGTTGGAAGGCGGCTGAGGCGGCTGAGGCGGCTGAGGCGGCTGAGGCGG 539  
 Db 480 GCAACTGTTGGAAGGCGGCTGAGGCGGCTGAGGCGGCTGAGGCGGCTGAGGCGG 539  
 Qy 483 GCAACTGTTGGAAGGCGGCTGAGGCGGCTGAGGCGGCTGAGGCGGCTGAGGCGG 542  
 Db 483 GCAACTGTTGGAAGGCGGCTGAGGCGGCTGAGGCGGCTGAGGCGGCTGAGGCGG 542  
 Qy 540 GGGATGTCGTCGAGGCGGCTGAGGCGGCTGAGGCGGCTGAGGCGGCTGAGGCGG 599  
 Db 540 GGGATGTCGTCGAGGCGGCTGAGGCGGCTGAGGCGGCTGAGGCGGCTGAGGCGG 599  
 Qy 543 GGGATGTCGTCGAGGCGGCTGAGGCGGCTGAGGCGGCTGAGGCGGCTGAGGCGG 602  
 Db 543 GGGATGTCGTCGAGGCGGCTGAGGCGGCTGAGGCGGCTGAGGCGGCTGAGGCGG 602  
 Qy 600 GTAAAGAGGCGGCGGCTGAGGCGGCTGAGGCGGCTGAGGCGGCTGAGGCGGCG 659  
 Db 600 GTAAAGAGGCGGCGGCTGAGGCGGCTGAGGCGGCTGAGGCGGCTGAGGCGGCG 659  
 Qy 603 GTAAAGAGGCGGCGGCTGAGGCGGCTGAGGCGGCTGAGGCGGCTGAGGCGGCG 660  
 Db 603 GTAAAGAGGCGGCGGCTGAGGCGGCTGAGGCGGCTGAGGCGGCTGAGGCGGCG 660  
 Qy 660 GCGAGTCCGCGGCTGAGGCGGCTGAGGCGGCTGAGGCGGCTGAGGCGGCTGAGG 719  
 Db 660 GCGAGTCCGCGGCTGAGGCGGCTGAGGCGGCTGAGGCGGCTGAGGCGGCTGAGG 719  
 Qy 661 GCCCGCCCTGAGGCTGAGGCGGCTGAGGCGGCTGAGGCGGCTGAGGCGGCTGAGG 720  
 Db 661 GCCCGCCCTGAGGCTGAGGCGGCTGAGGCGGCTGAGGCGGCTGAGGCGGCTGAGG 720  
 Qy 720 CGCTGCA---GAACAGCAGCTAGATTAACCTTAAGAAAGATTAATTTGACAGTAC 776  
 Db 720 CGCTGCA---GAACAGCAGCTAGATTAACCTTAAGAAAGATTAATTTGACAGTAC 776  
 Qy 721 CATGCCATGCTAATCCACAGCTGTTAAACCTTAAGAAAGATTAATTTGACAGTAC 780  
 Db 721 CATGCCATGCTAATCCACAGCTGTTAAACCTTAAGAAAGATTAATTTGACAGTAC 780  
 Qy 777 TTAAGATTAATCAATGCTAAATTTGAGGATGCTGTAATTCAGTCACTAATAGGCG 836  
 Db 777 TTAAGATTAATCAATGCTAAATTTGAGGATGCTGTAATTCAGTCACTAATAGGCG 836  
 Qy 781 TTCCTTGGGATCTTTGCAAAAGGCGGATGCTGAGGATGCTGAGGATGCTGAGGAT 839  
 Db 781 TTCCTTGGGATCTTTGCAAAAGGCGGATGCTGAGGATGCTGAGGATGCTGAGGAT 839

QY 837 GAATGGGATACGGGGCCCCCTCGAGGTGACGGATTCGATAGACTGATATGCAATTC 896  
| | | | |  
Db 840 GGAACAGGGCCATTCGCAATTCGATATTTGTTGATTAATGCTCTGATTCAGACGCTTC 899  
| | | | |  
QY 897 CTGACGCCCGGGGGAATCCACTAGTT----CTAGACGGCGCGCAACCGGGGTGAGCTCCA 952  
| | | | |  
Db 900 CATTCTTAATGTTGTGTCTAATTTTGAAGTTAACTTTGATTCATTTCTTTGTGTTGTCG 959  
| | | | |  
QY 953 GCTTTTGTCCCTTTAGTGAAGGTTAATTAATCCATGCGTCAATTTTAAAGCAGACTAT 1012  
| | | | |  
Db 960 CCATATGATATACATTTGTGTAGTTATGTTATTCATTTGTGTCAGAAATGTTTC 1019  
| | | | |  
QY 1013 CTTTCTAGGGTTAATCTAGCTCATCAGATATACGTGGGCTTTTTCGGCTAG 1072  
| | | | |  
Db 1020 CATCTCTTTAAATCAATTAACCTTTTAACTGATTTCTATTAACAAGGATACCTTCAA 1079  
| | | | |  
QY 1073 TCATCGCCCAAGCTGCGCTATCTGGGATCGGGAGAGAAAGCCGTCCTTTTCCCG 1132  
| | | | |  
Db 1080 ACTTGACTTCAAGCAGTGTCTGTAGTTCCCGTATCTTTGAAAATATATGTTCTTTCT 1139  
| | | | |  
QY 1133 CGAGGTGAAGCGGATGAAAAGATTTCGGAAGATGACTGTGCTGCATTTGACGTTGA 1192  
| | | | |  
Db 1140 GTACTTAACCTTCGGGCAATGCACTTTGAAAAAGTCATGCGTTCAATATGATCCGGGT 1199  
| | | | |  
QY 1193 GCGAAAACGACGTTTACATGATGATTGCGAAAGTGTGGCCATGACAGCCTTTTACG 1252  
| | | | |  
Db 1200 ATCTTGAAGACATTTGAACCACTAAGAGAAAGTGAACAAGTGTGGCCATGAAACAG 1259  
| | | | |  
QY 1253 TGAACGTGCTTCAAGGCCACCTGGGAATACCAATTCGTGCGGCTTTTCGGAACAGTT 1312  
| | | | |  
Db 1260 GTAGTTTTCAGTGTGCAATTAATTTAAGGTAAGTTTCCGATGTGTCATCACCTT 1319  
| | | | |  
QY 1313 CCGGATGGTCAAGCCGGAAGCGCATCAGAACCCGAACAATACCGGCGACAGCCGAACTG 1372  
| | | | |  
Db 1320 CACCTCTCCCACTCAAGAAATTTGTGCCATTTAACTAATCACTAATTAATCAACAAGAA 1379  
| | | | |  
QY 1373 CCGTGC-----GGTGTCAATTAATGACAGCGGTGCGGCTGCGGATAT 1418  
| | | | |  
Db 1380 TTGGGACAACCTCAGTGAAGGTTCTTCCTTTGCTAGACATTTCTTGGCCATGAAAT 1439  
| | | | |  
QY 1419 TACGTCAAGCAGAGCGGGTATCCGTGCTGATGCGCGAATATGACATGATATCCCGGT 1478  
| | | | |  
Db 1440 TCTGTCAGCCCGGGGATTCACATGATTCAGACGCGCGCACGCGGTGAACTCCAG 1499  
| | | | |  
QY 1479 GAGTTACCGGCGGG-----CGCGCTTGGCGTAACTGATGATAGCTGTT 1525  
| | | | |  
Db 1500 TTTTGTCCCTTAAGTAGAGGTTAATTCGAGCTTGGCGTATCATGTCATAGCTGTT 1559  
| | | | |  
QY 1526 CCGTGTGAATTTGTTATCCGCTCACAATTCACACAATACGAGCCGGAAACATTAAG 1585  
| | | | |  
Db 1560 CCGTGTGAATTTGTTATCCGCTCACAATTCACACAATACGAGCCGGAAACATTAAG 1619  
| | | | |  
QY 1586 TGTAAAGCTGGGGTGCCTAATGAGTGAAGCTAATCAATTAATCCGTTGGCTCACTG 1645  
| | | | |  
Db 1620 TGTAAAGCTGGGGTGCCTAATGAGTGAAGCTAATCAATTAATCCGTTGGCTCACTG 1679  
| | | | |  
QY 1646 CCGGCTTTTCCAGTCGGGAAACCTGTGCTGACAGCTCATTAATGATCGGCNAACGCGG 1705  
| | | | |  
Db 1680 CCGGCTTTTCCAGTCGGGAAACCTGTGCTGACAGCTCATTAATGATCGGCNAACGCGG 1739  
| | | | |  
QY 1706 GGGAGAGGGGTTTGGTATTTGGGCGCTTTCGCGTTCCTCGCTCACTGACTGCGTGGC 1765  
| | | | |  
Db 1740 GGGAGAGGGGTTTGGTATTTGGGCGCTTTCGCGTTCCTCGCTCACTGACTGCGTGGC 1799  
| | | | |  
QY 1766 TCGGTGTTTCGCTGCGGAGCGGATACGCTCACTCAAGGCGGTTAATGAGTTATCC 1825  
| | | | |  
Db 1800 TCGGTGTTTCGCTGCGGAGCGGATACGCTCACTCAAGGCGGTTAATGAGTTATCC 1859  
| | | | |  
QY 1826 AAGAAATCAGGGGATTAACGAGAGAAACATATGAGCAAAAGGCCAGCAAAAGGCCAGG 1885  
| | | | |  
Db 1860 AAGAAATCAGGGGATTAACGAGAGAAACATATGAGCAAAAGGCCAGCAAAAGGCCAGG 1919  
| | | | |

QY 1886 AACGTAATAAGGCGCGGTTGCTGSGGTTTTTCCATAGGTCGCGCCCCCTGACGACAT 1945  
| | | | |  
Db 1920 AACGTAATAAGGCGCGGTTGCTGSGGTTTTTCCATAGGTCGCGCCCCCTGACGACAT 1979  
| | | | |  
QY 1946 CACAAAATGACGCTCAAGTCAGAGGTGCGAACCACAGACATATTAAGATACAG 2005  
| | | | |  
Db 1980 CACAAAATGACGCTCAAGTCAGAGGTGCGAACCACAGACATATTAAGATACAG 2039  
| | | | |  
QY 2006 GCGTTTCCCCGTAAGGCTCCCTCGGCGCTCTCTGTTCCGACCTGCGCTTACCGGA 2065  
| | | | |  
Db 2040 GCGTTTCCCCGTAAGGCTCCCTCGGCGCTCTCTGTTCCGACCTGCGCTTACCGGA 2099  
| | | | |  
QY 2066 TACCTGTCCGCTTTCTCCCTTCGGAACGCTGCGCTTCTCATAGTCAGCCTGTAG 2125  
| | | | |  
Db 2100 TACCTGTCCGCTTTCTCCCTTCGGAACGCTGCGCTTCTCATAGTCAGCCTGTAG 2159  
| | | | |  
QY 2126 TACTCATGTTGCGTGAAGTCTGTCCTCAAGCTGGGCTGTGACGAACCCCGTT 2185  
| | | | |  
Db 2160 TACTCATGTTGCGTGAAGTCTGTCCTCAAGCTGGGCTGTGACGAACCCCGTT 2219  
| | | | |  
QY 2186 CAGCCGACCGGCTGCGCTTATCCGGTAACTATGCTTTGAGTCCAACCCGTTAAGAC 2245  
| | | | |  
Db 2220 CAGCCGACCGGCTGCGCTTATCCGGTAACTATGCTTTGAGTCCAACCCGTTAAGAC 2279  
| | | | |  
QY 2246 GACTTATCCGCACTGGCAGACCCA CTGGTAA CAGAAATTAAGCAGACGATATGAGC 2305  
| | | | |  
Db 2280 GACTTATCCGCACTGGCAGACCCA CTGGTAA CAGAAATTAAGCAGACGATATGAGC 2339  
| | | | |  
QY 2306 GGTGCTACAGAGTTCTTAAGTGTGGGCTTAATCCGCTTACATGAAGACAGTATTT 2365  
| | | | |  
Db 2340 GGTGCTACAGAGTTCTTAAGTGTGGGCTTAATCCGCTTACATGAAGACAGTATTT 2399  
| | | | |  
QY 2366 GGTATCTGGGCTCTGCTGAAGCAGTTACTTTCGGAATAAGTTGTAAGCTTGAATCC 2425  
| | | | |  
Db 2400 GGTATCTGGGCTCTGCTGAAGCAGTTACTTTCGGAATAAGTTGTAAGCTTGAATCC 2459  
| | | | |  
QY 2426 GGGAAACAAACCAACGCTGTGAGCGGTGTTTTTTGTTTCAAGACAGCAATTAACGCG 2485  
| | | | |  
Db 2460 GGGAAACAAACCAACGCTGTGAGCGGTGTTTTTTGTTTCAAGACAGCAATTAACGCG 2519  
| | | | |  
QY 2486 AGAAAAAAGGATCTCAAGAGATCTTGTATCTTTTCTAGGGGCTGACGCTCACTG 2545  
| | | | |  
Db 2520 AGAAAAAAGGATCTCAAGAGATCTTGTATCTTTTCTAGGGGCTGACGCTCACTG 2579  
| | | | |  
QY 2546 AAGAAAACTACGTTAAGGATTTTGTCTAGATTAATCAAAAAGATCTTCACTAG 2605  
| | | | |  
Db 2580 AAGAAAACTACGTTAAGGATTTTGTCTAGATTAATCAAAAAGATCTTCACTAG 2639  
| | | | |  
QY 2606 ATCTTTTAATTTAAAAATGAAGTTTAAATCATCTTAAGTATATATGATTAACCTGG 2665  
| | | | |  
Db 2640 ATCTTTTAATTTAAAAATGAAGTTTAAATCAATCTTAAGTATATATGATTAACCTGG 2699  
| | | | |  
QY 2666 TCTGACGTTTCAATGCTTAATCAGTGAAGCACTTCTCAGGATCTGTATTTGCT 2725  
| | | | |  
Db 2700 TCTGACGTTTCAATGCTTAATCAGTGAAGCACTTCTCAGGATCTGTATTTGCT 2759  
| | | | |  
QY 2726 TCAATCCATAGTGGCTGATCTCCGCTGTGTATTAATCTAGATTAACGGAAGGCTTACA 2785  
| | | | |  
Db 2760 TCAATCCATAGTGGCTGATCTCCGCTGTGTATTAATCTAGATTAACGGAAGGCTTACA 2819  
| | | | |  
QY 2786 TCTGAGCCCAAGTGTGAATGATACGCGAGACCAACGCTCACCGGCTCCAGATTAATCA 2845  
| | | | |  
Db 2820 TCTGAGCCCAAGTGTGAATGATACGCGAGACCAACGCTCACCGGCTCCAGATTAATCA 2879  
| | | | |  
QY 2846 GCAATTAACCAAGCCAGCGGAGGCGGAGCGGAGAAAGTGTCTGCAACTTATCCGCG 2905  
| | | | |  
Db 2880 GCAATTAACCAAGCCAGCGGAGGCGGAGCGGAGAAAGTGTCTGCAACTTATCCGCG 2939  
| | | | |  
QY 2906 TCAATCAAGCTTATTAATGTTGTCGCGGAAGCTAAGATAGTATCCGCAAGTTAATAGT 2965  
| | | | |  
Db 2940 TCAATCAAGCTTATTAATGTTGTCGCGGAAGCTAAGATAGTATCCGCAAGTTAATAGT 2999  
| | | | |  
QY 2966 TTGCGCAACGTTGTGCTACAGGCAATGCTGTGTGTCACGCTGCTGTTGTGATG 3025  
| | | | |

Db 3000 TTGGGCAAGTGTGGCCATGCTACAGGCAATCGTGTGTCACGCTCGTGTGTATG 3059  
 Qy 3026 GCTTCATTCAGCTCCGCTTCCCAACGATCAAGCGAGTTACATGATCCCAATGTTGTGC 3085  
 Db 3060 GCTTCATTCAGCTCCGCTTCCCAACGATCAAGCGAGTTACATGATCCCAATGTTGTGC 3119  
 Qy 3086 AAAAAGGGGTAGCTCCTCGGTCCGATCCGATCCGTTGTCAGAAAGTAAGTTGGCCGACTG 3145  
 Db 3120 AAAAAGGGGTAGCTCCTCGGTCCGATCCGATCCGTTGTCAGAAAGTAAGTTGGCCGACTG 3179  
 Qy 3146 TTATCACTCATGTTATGAGCACTGATCAATATCTCTTAATGTCATGCAATCCGTAAGA 3205  
 Db 3180 TTATCACTCATGTTATGAGCACTGATCAATATCTCTTAATGTCATGCAATCCGTAAGA 3239  
 Qy 3206 TGCCTTTCTGTCAGCTGTGAGTACTCAACCAAGTCAATCTGAGAAATGTAATGCGGCGA 3265  
 Db 3240 TGCCTTTCTGTCAGCTGTGAGTACTCAACCAAGTCAATCTGAGAAATGTAATGCGGCGA 3299  
 Qy 3266 CCGAGTGTCTGTGCGCGCGCTCAATAGGGAATATACCGGCGCATAGCAAGAACTTTA 3325  
 Db 3300 CCGAGTGTCTGTGCGCGCGCTCAATAGGGAATATACCGGCGCATAGCAAGAACTTTA 3359  
 Qy 3326 AAAGTGTCTCATTTGGAAGAAAGTCTTCCGCGCGAAGAACTCTCAAGGATCTTACCGCTG 3385  
 Db 3360 AAAGTGTCTCATTTGGAAGAAAGTCTTCCGCGCGAAGAACTCTCAAGGATCTTACCGCTG 3419  
 Qy 3386 TTGAGATCACTGTCATGATGTAACCACTGTCGCAATCTGATCTTCAAGTCTTTTACT 3445  
 Db 3420 TTGAGATCACTGTCATGATGTAACCACTGTCGCAATCTGATCTTCAAGTCTTTTACT 3479  
 Qy 3446 TTCCACGAGCGTTTCTGGGTGAGCAAAAAAGAGGCAAAATGCCGCAAAAAAGGGAATA 3505  
 Db 3480 TTCCACGAGCGTTTCTGGGTGAGCAAAAAAGAGGCAAAATGCCGCAAAAAAGGGAATA 3539  
 Qy 3506 AGGCGCAACGGAAGATGTTGAATCTCATACTCTTCTTTTCAATATTTATTTGAAGCAAT 3565  
 Db 3540 AGGCGCAACGGAAGATGTTGAATCTCATACTCTTCTTTTCAATATTTATTTGAAGCAAT 3599  
 Qy 3566 TATCAGAGGTTATGTCATGATGAGCGGATACATATTTGAATGATTTAGAAAAATTAACA 3625  
 Db 3600 TATCAGAGGTTATGTCATGATGAGCGGATACATATTTGAATGATTTAGAAAAATTAACA 3659  
 Qy 3626 ATAGGGGTTCGCGCAATTTCCCGGAAAAAGTGCAC 3662  
 Db 3660 ATAGGGGTTCGCGCAATTTCCCGGAAAAAGTGCAC 3696

RESULT 14

AAD27066  
 ID AAD27066 standard; DNA; 4001 BP.  
 AC AAD27066;  
 XX  
 DT 09-APR-2002 (first entry)  
 DE Plasmid pGN59A DNA.  
 XX  
 KM Vector construct; RNA inhibition; RNAi; gene expression control;  
 KM pGN49A plasmid; ds.  
 XX  
 OS Unidentified.  
 XX  
 XX WO20018121-A1.  
 FN  
 XX 22-NOV-2001.  
 PD  
 XX 18-MAY-2001; 2001WO-IB001068.  
 PF  
 XX 19-MAY-2000; 2000GB-00012233.  
 PR  
 XX (DEVG-) DEVGEN NV.  
 PA  
 XX

PI Placitnick G, Renard J, Bogaert T;  
 XX  
 DR WPI; 2002-121984/16.  
 XX  
 PT A new DNA vector construct containing opposable promoter and terminator  
 PT sequences flanking a cloning site are useful for the expression of double  
 PT stranded RNA useful for inhibition of RNA in gene expression control.  
 XX  
 ES Claim 24; Fig 15; 75bp; English.  
 CC  
 CC The present invention relates to improved vector constructs comprising  
 CC two promoters in opposite orientation to each other, an inter-promoter  
 CC region downstream of the 3' end of both promoters, a cloning site in the  
 CC inter-promoter region and a transcription terminator downstream of the 3',  
 CC end of the first promoter and the cloning site and operably linked to the  
 CC first promoter. The constructs of the invention and the bacteria  
 CC harboring the constructs are used to produce double stranded RNA for RNA  
 CC inhibition (RNAi) and can be used as a tool for controlling gene  
 CC expression. The present sequence is pGN59A plasmid DNA  
 XX  
 SQ Sequence 4001 BP; 1034 A; 967 C; 1030 G; 970 T; 0 U; 0 Other;  
 Query Match 71.3%; Score 2609.4; DB 6; Length 4001;  
 Best Local Similarity 83.1%; Pred. No. 0;  
 Matches 3069; Conservative 0; Mismatches 591; Indels 35; Gaps 7;  
 Qy 1 CTAATGTTAGAGCTTAAATATTTTGTAAATTTGCGTTAAATTTTGTAAATCAGCTC 60  
 Db 63 CGAAATGTTAAAGCTTAAATATTTTGTAAATTTGCGTTAAATTTTGTAAATCAGCTC 122  
 Qy 61 ATTTTAAACCAATAGCCGCAAAATCCCTTAAATCAAAAGAAATGACCGA 120  
 Db 123 ATTTTAAACCAATAGCCGCAAAATCCCTTAAATCAAAAGAAATGACCGA 182  
 Qy 121 GATAGGGTGAAGTGTGTTCCAGTTTGAACCAAGTCCACTATTTAAAGAACTGAGACTC 180  
 Db 183 GATAGGGTGAAGTGTGTTCCAGTTTGAACCAAGTCCACTATTTAAAGAACTGAGACTC 242  
 Qy 181 CAAGTCAAAGGGGCAAAAAACGTCATCAGGGGAGTGGCCACTAGCTGACATCAACC 240  
 Db 243 CAAGTCAAAGGGGCAAAAAACGTCATCAGGGGAGTGGCCACTAGCTGACATCAACC 302  
 Qy 241 CTAATCAAGTTTGTGGGTGTCAGTCCGTTAAAGCACTAAATGGAACCTTAAGGAG 300  
 Db 303 CAAATCAAGTTTGTGGGTGTCAGTCCGTTAAAGCTCTAAATGGAACCTTAAGGAG 362  
 Qy 301 CCCCCGATTTAGACTTGAACGGGGAAGCCGGCAACGTGGCGAAGAAAGGAAGAA 360  
 Db 363 CCCCCGATTTAGACTTGAACGGGGAAGCCGGCAACGTGGCGAAGAAAGGAAGAA 422  
 Qy 361 AGCGAAAGAGCGGGGCTAGGGGCTGAGCAAGTGAAGCGTCAACGCTGCCCTAACAC 420  
 Db 423 AGCGAAAGAGCGGGGCTAGGGGCTGAGCAAGTGAAGCGTCAACGCTGCCCTAACAC 482  
 Qy 421 CACACCGCGCGCTTAAATGCGCGCTCAACAGGGGCGCTCCATTCGCTACAGCTGCG 480  
 Db 483 CACACCGCGCGCTTAAATGCGCGCTCAACAGGGGCGCT -CCATTCGCTACAGCTGCG 541  
 Qy 481 CAACTGTGGAAGAGGCGATCGGTGCGGCGCTTCTGCTATTACGCAAGCTGGGAAAG 540  
 Db 542 CAACTGTGGAAGAGGCGATCGGTGCGGCGCTTCTGCTATTACGCAAGCTGGGAAAG 601  
 Qy 541 GGGATGTGTCGAAGGCGATTAAGTTGGTAAACGCCAGGTTTTTCCAGTCAAGCACTG 600  
 Db 602 GGGATGTGTCGAAGGCGATTAAGTTGGTAAACGCCAGGTTTTTCCAGTCAAGCACTG 661  
 Qy 601 TAAACGACGCGCCAGTGAAGCGCGCT-----CGTTCATTCAGCTTTTGA 646  
 Db 662 TAAACGACGCGCCAGTGAATTTGAATGCACTCACTAAGGGCGAATTCAGCTCGGTAC 721  
 Qy 647 CCCGTGAGAGACGGGCGAGCTCGCGGTGCAATGTGTTTACAGCGTATGAGCAATG 706  
 Db 722 CCCGGGATCCTCTTAAGATCCTCTGACCTGAGATCATTTGTGCTGAAAGATCTGAT 781



QY 707 AAGATGCTCGACAGCTGCGAACAACGACGACTAGT-----TAACTTGAAGAAATGA 759  
Db 782 CCGGCTTACTAATAAGCCAGATATACATATGCTATTGGCGGCTGATTTTGGCGATATA 841  
QY 760 TCATATTGTGACAGTACGTTAAAGATATCATGCTATAATTG-----ACGATGGGAT 812  
Db 842 GAATATATATGATATATATCCGAAAGTATGTCAAAAAGAGGTGCTATGAAAGCAGCG 901  
QY 813 CTGTAATACGACTACTATAGGGGGAATTGGGTACCGGGCCCCCTCGAGGTGACGGT 872  
Db 902 TATTACGATCAGTTGACAGCGACAGCTATCAGTTGCTCAAGGCATATATGATGCAAT 961  
QY 873 ATCGATAGCTTGAATATCGAATTCTGAGCCCGGGGAGTCCAATGTTCTAGAGCGGC 932  
Db 962 ATCTCGGCTGTGTAAGCACAACATGCAAGATGAAGCCCGTGTCTGCGTCCGAACGC 1021  
QY 933 GCCACCGGGTGAGCTCCAGCTTTGTTCCCTTTAGTGGGTTAATTAGATCCCATGC 992  
Db 1022 TGGAAAGCGGAAATCAGGAAGGGATGGCTGAGGTGGCCGGTTTATGAAATGAACGAC 1081  
QY 993 GTCAATTTTACGACGACTA---TCTTTAGGTTATCTAGCTGCATCAGATCATATC 1049  
Db 1082 TCTTTTGCTGACGAGAACAGGACTGTGTAATGCAGTTTAAAGTTTACACTATTAAG 1141  
QY 1050 GTGCGGCTTTTTCGAGCTCAGTATCGCCCAAGCTGGCGCTATCTGGGCAATCGGGAG 1109  
Db 1142 AGAGAGCCGTATCGTCTGTTGTGGATGTACAGATATATTTAGACAGCCCGGGCG 1201  
QY 1110 GAAGAGCCCGTGGCTTTCCCGCAGAGTTGAAGCGCATGGAAGAAGTTTCCGAGGAT 1169  
Db 1202 ACGGATGTGATCCCTCGGCC-AGTGAACGCTCTTTAAGCGATPAAGTCTCCGCTGAC 1260  
QY 1170 GACTGCTGCT--CATTGAAGTTGAGCAAAAACGACGTTTACATGATGATTTGGGAG 1227  
Db 1261 TTATCCCGTGTGATATCGGGAGTAAAGCTGGCGATATACACGCAATATGCGCA 1320  
QY 1228 GTGCGCATGACCGCTTTAAAGGTGAATGTTCTGACGCGCACTGGGATACAGTT 1287  
Db 1321 GTGTCCCGCTCTCCGTTATCGGGAAAGAGTGGCTGATCTCAGCAACGGAATATGCA 1380  
QY 1288 CGTCCGGGCTTTTCGGAACAAGTCCGATGTCAGCCGGAAGCGATACGACCCGA 1347  
Db 1381 TCAAAAAGCCATTAACCTGATGTTCTGGGAAATATTAATGTCAAGGCTCTCTTATAC 1440  
QY 1348 ACAATACCGGCGAACAAGCCGAACCTGCGGTGAGATTAATGACAGCGGTGCG 1407  
Db 1441 GCTTTTCAGCAACATGATCTTCAGGGATCTTTCATACCTACCAAGTTTGGCCCTGAG 1500  
QY 1408 CGTGGGATATTAAGTACGAGGAGCGGGTATCTGGCTGGATGCGCGCAAAATGACAT 1467  
Db 1501 GTCCGCGCCGGAATCTTAGAGTCGAAAGCTTCTGCGCTTATAGTGAATCGTATTAAG 1560  
QY 1468 GGAATACCCCGATTAACCGGGGGGCGCTTGGGTATCATGTGTCACTAGCTGTTCC 1527  
Db 1561 CTGAGATTTCTATAGTGTCACTAAATAGCTTGGGTATATCATGCTATAGCTGTTCC 1620  
QY 1528 TGTGTGAATTTGTATCCGCTCACAATTCACACAACAATGACAGCCGGAAGCATTAAGT 1587  
Db 1621 TGTGTGAATTTGTATCCGCTCACAATTCACACAACAATGACAGCCGGAAGCATTAAGT 1680  
QY 1588 TAAAGCTGGGGTCCCTAATAGTGAAGTAACTCACTAATTAATGCTTGGCTGACTGCC 1647  
Db 1681 TAAAGCTGGGGTCCCTAATAGTGAAGTAACTCACTAATTAATGCTTGGCTGACTGCC 1740  
QY 1648 CGCTTTCAGTCCGGAACCTGTGCGCAAGTGCATTAATGATCGGCCAACCGCGGG 1707  
Db 1741 CGCTTTCAGTCCGGAACCTGTGCGCAAGTGCATTAATGATCGGCCAACCGCGGG 1800  
QY 1708 GAGAGCGGTTTGCATTTGGCGCTCTTTCGCTTCTGCTCACTGATGCTGCGCTC 1767  
Db 1801 GAGAGCGGTTTGCATTTGGCGCTCTTTCGCTTCTGCTCACTGATGCTGCGCTC 1860

QY 1768 GGTGCTTCGCTGCGGAGAGCGGTATCAGTCACTCAAGCGGGTATATACGTTATCCAC 1827  
Db 1861 GGTGCTTCGCTGCGGAGAGCGGTATCAGTCACTCAAGCGGGTATATACGTTATCCAC 1920  
QY 1828 AGAATCAGGGGATTAACCAAGAAAGAAATGTGACAAAAGGCGCAAAAAGGCGAGAA 1887  
Db 1921 AGAATCAGGGGATTAACCAAGAAAGAAATGTGACAAAAGGCGCAAAAAGGCGAGAA 1980  
QY 1888 CCGTAAAGCGCGCGTGTGCGCTTTTTCATAGAGTCCGCGCCCTGACAGAGATCA 1947  
Db 1981 CCGTAAAGCGCGCGTGTGCGCTTTTTCATAGAGTCCGCGCCCTGACAGAGATCA 2040  
QY 1948 CAAAATTCGACGCTCAAGTATAGAGTGGGAAACCCGACAGACTATAAAGATACAGGC 2007  
Db 2041 CAAAATTCGACGCTCAAGTATAGAGTGGGAAACCCGACAGACTATAAAGATACAGGC 2100  
QY 2008 GTTTCGCCCTGAAAGCTCCCTGCGGCTCTCTGTTCGACCCCTGCGCTTACCGGAT 2067  
Db 2101 GTTTCGCCCTGAAAGCTCCCTGCGGCTCTCTGTTCGACCCCTGCGCTTACCGGAT 2160  
QY 2068 CCGTCCGCGCTTCTCTCCCTTCCGGAAGCGTGGCGTTTCTCATAGCTCAGCTGATGTA 2127  
Db 2161 CCGTCCGCGCTTCTCTCCCTTCCGGAAGCGTGGCGCTTCTCATAGCTCAGCTGATGTA 2220  
QY 2128 TCTCAGTTGAGTGTAGTGTGTTGCTCCAGAGCTGGGCTGTGTGACGAACCCCGTTCA 2187  
Db 2221 TCTCAGTTGAGTGTAGTGTGTTGCTCCAGAGCTGGGCTGTGTGACGAACCCCGTTCA 2280  
QY 2188 GCCCAGCGCTGCGCTTATCCGGAATCTATGTTGATGTCGAACCGGTAAGACGA 2247  
Db 2281 GCCCAGCGCTGCGCTTATCCGGAATCTATGTTGATGTCGAACCGGTAAGACGA 2340  
QY 2248 CTATTCGCACTGAGCAGAGCACTGTGAACAGATTAAGCAGACCGAGTATGTAAGCGG 2307  
Db 2341 CTATTCGCACTGAGCAGAGCACTGTGAACAGATTAAGCAGACCGAGTATGTAAGCGG 2400  
QY 2308 TGTCTAAGATTTTGAAGTGTGGCTTAACGAGCTACACTGAGAGCAAGTATTTGG 2367  
Db 2401 TGTCTAAGATTTTGAAGTGTGGCTTAACGAGCTACACTGAGAGCAAGTATTTGG 2460  
QY 2368 TATCTGCGCTGCTGGAAGCAGTTACCTCGAAAAAGAGTTGTAAGTCTTGTATCCG 2427  
Db 2461 TATCTGCGCTGCTGGAAGCAGTTACCTCGAAAAAGAGTTGTAAGTCTTGTATCCG 2520  
QY 2428 CAACCAACCAACCGCTGTAGCGGTGTTTTTGTGTGCAACAGAGATTAAGCGGAG 2487  
Db 2521 CAACCAACCAACCGCTGTAGCGGTGTTTTTGTGTGCAACAGAGATTAAGCGGAG 2580  
QY 2488 AAAAAAGATCTCAAGAAAGATCTTGTATCTTTTCTAGGGGTCTGACGCTCAGTGAA 2547  
Db 2581 AAAAAAGATCTCAAGAAAGATCTTGTATCTTTTCTAGGGGTCTGACGCTCAGTGAA 2640  
QY 2548 CGAAACTCAAGTTAAGGATTTTGTGTCATGATTAATCAAAAAGGATCTTCACTAGAT 2607  
Db 2641 CGAAACTCAAGTTAAGGATTTTGTGTCATGATTAATCAAAAAGGATCTTCACTAGAT 2700  
QY 2608 CTTTTTAAATTAATTAAGATTTTAATCAATCTTAAGATATATGATTAATCTTGTGTC 2667  
Db 2701 CTTTTTAAATTAATTAAGATTTTAATCAATCTTAAGATATATGATTAATCTTGTGTC 2760  
QY 2668 TGAAGTTTACATGCTTATCAGTGAAGGACCTTATCAGAGATCTGTCTATTTTCGTT 2727  
Db 2761 TGAAGTTTACATGCTTATCAGTGAAGGACCTTATCAGAGATCTGTCTATTTTCGTT 2820  
QY 2728 ATCCATAGTTTGTGCTGACTCCCGTGTGTGATTAATCAATCAGTACGAGGAGCTTACATC 2787  
Db 2821 ATCCATAGTTTGTGCTGACTCCCGTGTGTGATTAATCAATCAGTACGAGGAGCTTACATC 2880  
QY 2788 TGGCCCAAGTGTGATATATACGAGAGCAACAGCTCAACGGCTCAGATTTTATCAGC 2847  
Db 2881 TGGCCCAAGTGTGATATATACGAGAGCAACAGCTCAACGGCTCAGATTTTATCAGC 2940  
QY 2848 AATTAACCAAGCCAGCGGAGCGGAGCGCAAGATGTGTCTTCAACTTTATCCGCTC 2907





QY 601 TAAACGACGCGCACTGAGCGCGCTGTCATTCACGTTTGGAACCCGTGAGAGCGG 660  
Db 662 TAAACGACGCGCACTGAGCGCGCTGTCATTCACGTTTGGAACCCGTGAGAGAGCGG 677  
QY 661 GCGACTTCGCGGTCAAAATGTTTATACAGCGTGTATGAGACAGATGAAATGCTCGACAC 720  
Db 678 ----- 677  
QY 721 GCTGACAGACGCGAGCTAGATTACCTTAGAAAATATCATATTTGTGACGTAA 780  
Db 678 ----- 677  
QY 781 AGATTAATCATGCGTAAATTAATGACATGCGATTCGTATATACATCTATATGCGCAAT 840  
Db 678 ----- GAATTTAATTAACATCTATATGCGCGAAT 708  
QY 841 TTGGGTACCGGCGCCCCCTGAGGTGACCGGTATCGATTAAGCTTGATATCGAATTC-CT 898  
Db 709 TCAAAAAACCCCTCAAGACCCGTTTAGAGGCCCAAGGGGTTATGCTAGTGAATTTCTGCA 768  
QY 899 GCAGCCCGGCGGATCCACTAGTTCTAGAGCGGCGGCACCGCGGTGAGCTCCAGC--T 955  
Db 769 GGGTACCCGGGATCTCTAGAGATCCCTGACCTCGAGATCATTTGTGCGCGGAT 828  
QY 956 TTTGTTCCCTTTAGTGAAGGTTAATTGAATCCATGCGCAATTTTACGACAGATATCTT 1015  
Db 829 TCTTTATCACTGAATAGTTGTGTGACATATATATGTTTATCAGTGATTAAGTCAAGCAT 888  
QY 1016 TCTAGGGTTAATCTAGCTGCATCAGATCATATCGTCCGGCTTTTTCGCGCTCAGTCA 1075  
Db 889 GACAAAGTTGACCGG-AAATACATGATCCGTGCGGCCCTGGAATGTTGAACAGAGTTCG 947  
QY 1076 TCGCCCAAGCTGCGCTATCTGGGCATCGGAGAGAGAACCCGCTGCTTTTCCCGCGCA 1135  
Db 948 GGTAGACGCTCTGAACGACAACTGGCGGAACGGTTGGGGTGCAGACGCGCGC 1007  
QY 1136 GGTAAAGCGGATGAGAAAGATTGCCGAGATGACTGCTGTCATTTAGCGTTAGCG 1195  
Db 1008 TTTACTGCACTTCAGGAAACAGCGGCGCTGCTGACGCACTGGCCGAAGCATGCTGG 1067  
QY 1196 AAAACGACGTTTATCCATGATGATTCGGAAGGTGTGGCAATGACGCTTTAAACGTGA 1255  
Db 1068 CGGAATCATACCTTCGCTGCGAGGCGAGAGACGATGCGCTCATTTCTGATTCG 1127  
QY 1256 ACTGTTGCTGAGGCACTCGGATACCAATTCGTGCGGCTTTTCCGACACATTCGCG 1315  
Db 1128 GAA-----TCCGCAAGCTTCAGGCGCGCTGCTGCTTACCGCAGACCAATGATCT 1181  
QY 1316 GATGTCAGCCCGAAGGCAATCAGCAACCGAACTAACCGGGAACGCGGAACTGCGG 1375  
Db 1182 CGAGGATCTTCCATCTTCAAGTTCTGCGCTGCAAGTCCGCGCGGACTCTCTAGA 1241  
QY 1376 TGCAGGTGCAATTATGACAGCGGTGCGGCTGGGATATTAAGTCAAGGAGAGCG 1435  
Db 1242 CGGTAAGCTTATCAGCAATAACCTTGGGCGCTTAAACGGGCTTGAAGGGTTTTTGG 1301  
QY 1436 GTATCTGCGTGAATGCGGAAATGACATGACATACCCCGATTAACCGCGGCGG 1495  
Db 1302 AGCTTCTCGCCTATATGTAGTGTATTCAGTTTGAATTCATATGATCTACCTTAAT 1361  
QY 1496 CGCTGGCGTAAATCATAGCTCATAGCTTTTCTGTGTGAATTTGTTATCCGTCACAAAT 1555  
Db 1362 AGTTTGGCGTAATCATAGCTCATAGCTTTTCTGTGTGAATTTGTTATCCGTCACAAAT 1421  
QY 1556 CCAACAAATATCAGACCGGAGCAATAAAGTAAAGCTGGGGTCTTAATGATGAGC 1615  
Db 1422 CCAACAAATATCAGACCGGAGCAATAAAGTAAAGCTGGGGTCTTAATGATGAGC 1481  
QY 1616 TAACTCAGTAAATGCGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1675  
Db 1482 TAACTCAGTAAATGCGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1541

QY 1676 CAGCTGCAATTAATATCGGCAACGCGCGGAGAGAGCGGTTTGGATATTTGGGCGCTCT 1735  
Db 1542 CAGCTGCAATTAATATCGGCAACGCGCGGAGAGAGCGGTTTGGATATTTGGGCGCTCT 1601  
QY 1736 TCCGCTTCTCGCTCACTGACTGCGCGGCTCGGCTGTTGGCTGCGGCGAGCGGTATCA 1795  
Db 1602 TCCGCTTCTCGCTCACTGACTGCGCGGCTCGGCTGTTGGCTGCGGCGAGCGGTATCA 1661  
QY 1796 GCTCACTCAAAAGCGGTAAATACGTTATCCAAGAAATCAGGGGATTAACGAGAAAGAC 1855  
Db 1662 GCTCACTCAAAAGCGGTAAATACGTTATCCAAGAAATCAGGGGATTAACGAGAAAGAC 1721  
QY 1856 ATGTGACCAAAAGCGGCAAAAGCGGCAAGAACCTTAAAGCGCGGTTGCTGCGCTT 1915  
Db 1722 ATGTGACCAAAAGCGGCAAAAGCGGCAAGAACCTTAAAGCGCGGTTGCTGCGCTT 1781  
QY 1916 TTCCATAGGCTCGGCCCTCGAGCATCACAAAATCGAGCTCAAGTCAAGGTG 1975  
Db 1782 TTCCATAGGCTCGGCCCTCGAGCATCACAAAATCGAGCTCAAGTCAAGGTG 1841  
QY 1976 CGAAACCGGACAGGATTAATTAAGATTCAGGCGTTTCCCTGGAAGCTCCCTGTCGCG 2035  
Db 1842 CGAAACCGGACAGGATTAATTAAGATTCAGGCGTTTCCCTGGAAGCTCCCTGTCGCG 1901  
QY 2036 TCTCTGTTCCGACCTCGCGCTTACCGGATACCTGTCGCGCTTCTCCCTGCGGAAC 2095  
Db 1902 TCTCTGTTCCGACCTCGCGCTTACCGGATACCTGTCGCGCTTCTCCCTGCGGAAC 1961  
QY 2096 GTGGGCTTTCATATGCTCACGCTGTATGATCTCAATTTGGATAGTGTGCTGCTGCT 2155  
Db 1962 GTGGGCTTTCATATGCTCACGCTGTATGATCTCAATTTGGATAGTGTGCTGCTGCT 2021  
QY 2156 AAGCTGGCTGTGTCAGAAACCCCGCTTACGCGGCAAGCGCTGCGCTTATCCGCTAAC 2215  
Db 2022 AAGCTGGCTGTGTCAGAAACCCCGCTTACGCGGCAAGCGCTGCGCTTATCCGCTAAC 2081  
QY 2216 TATGCTTGTGATCCAAACCGGTAAGACAGACTTATCGCACTGAGAGCACTGCT 2275  
Db 2082 TATGCTTGTGATCCAAACCGGTAAGACAGACTTATCGCACTGAGAGCACTGCT 2141  
QY 2276 AACAGATTTAGCAGAGGATATGTAGGCGGTGCTACAGATTTCTTGAATGTGCGCT 2335  
Db 2142 AACAGATTTAGCAGAGGATATGTAGGCGGTGCTACAGATTTCTTGAATGTGCGCT 2201  
QY 2336 AACTACGCTTACTTAAAGCAAGTATTTGTATCTGCGCTGCTGCTGCTGCTGCTGCT 2395  
Db 2202 AACTACGCTTACTTAAAGCAAGTATTTGTATCTGCGCTGCTGCTGCTGCTGCTGCT 2261  
QY 2396 TTCCGAAAAAGATGTGATCTTGAATCCGGAACCAACCGCTGCTGCTGCTGCTGCT 2455  
Db 2262 TTCCGAAAAAGATGTGATCTTGAATCCGGAACCAACCGCTGCTGCTGCTGCTGCT 2321  
QY 2456 TTTTGTGTTGCAAGCAGATTTACGCGCAGAAAAAAGATCTCAAGAAATCTTTT 2515  
Db 2322 TTTTGTGTTGCAAGCAGATTTACGCGCAGAAAAAAGATCTCAAGAAATCTTTT 2381  
QY 2516 ATCTTTTCTTACGCGGCTGACGCTGAGGGAACGAAATCTCAGTTAAGGATTTTGTG 2575  
Db 2382 ATCTTTTCTTACGCGGCTGACGCTGAGGGAACGAAATCTCAGTTAAGGATTTTGTG 2441  
QY 2576 ATAGATTTATCAAAAGATCTTCACTAGATCTTTTAAATTAATAAGATTTTAA 2635  
Db 2442 ATAGATTTATCAAAAGATCTTCACTAGATCTTTTAAATTAATAAGATTTTAA 2501  
QY 2636 TCAATTAAGATTAATTAAGATTTAAGATTTGCTGACGTTTACCAATGCTTATCAGTGA 2695  
Db 2502 TCAATTAAGATTAATTAAGATTTAAGATTTGCTGACGTTTACCAATGCTTATCAGTGA 2561  
QY 2696 GACCTATCTCAGCGATGCTATTTGCTCATCATAGTTGCTGACCTCCCGTGTG 2755  
Db 2562 GACCTATCTCAGCGATGCTATTTGCTCATCATAGTTGCTGACCTCCCGTGTG 2621  
QY 2756 TAGATTAATCAGATTCGGAAGGCGCTTACATCTGCCCCAGTGTGATGATACCGCA 2815

```
Db 2622 TAGATACTAGATACGGAGGGCTTACCATCTGGCCCGCAGTGGCTGCAATGATACCGCA 2681
QY 2816 GACCCAGGCTACCGGCTCCGATTTATCAGCAATAAACAGCCAGCCGGAAGGCGGAG 2875
Db 2682 GACCCAGGCTACCGGCTCCGATTTATCAGCAATAAACAGCCAGCCGGAAGGCGGAG 2741
QY 2876 CGCAGAGTGGTCCCTGCAACTTTATCCGCTCCATCCAGTCAATTAATTGTGGCGGGA 2935
Db 2742 CGCAGAGTGGTCCCTGCAACTTTATCCGCTCCATCCAGTCAATTAATTGTGGCGGGA 2801
QY 2936 GCTAGAGTAAGTAAGTTCGCACTTAATAGTTGCGCAAGTTGTGCAATGCTACAGGC 2995
Db 2802 GCTAGAGTAAGTAAGTTCGCACTTAATAGTTGCGCAAGTTGTGCAATGCTACAGGC 2861
QY 2996 ATCGTGTGTACGCTGCTGTTGGTATGCTTTCATTCAGCTCCGTTCCCAAGATCA 3055
Db 2862 ATCGTGTGTACGCTGCTGTTGGTATGCTTTCATTCAGCTCCGTTCCCAAGATCA 2921
QY 3056 AGGAGATTACATGATCCCGCATGTTGCAAAAAAGGGTTAGCTCCTTGGTCCGCG 3115
Db 2922 AGGAGATTACATGATCCCGCATGTTGCAAAAAAGGGTTAGCTCCTTGGTCCGCG 2981
QY 3116 ATCGTTGCAAGTAAGTGGCCGCAAGTTATCACTCATGTTATGGCAGCACTGCAT 3175
Db 2982 ATCGTTGCAAGTAAGTGGCCGCAAGTTATCACTCATGTTATGGCAGCACTGCAT 3041
QY 3176 AATTCTCTTACTGTGATGCAATCCGTAAGATGCTTTCTGTGACTGTGAGTACTCAAC 3235
Db 3042 AATTCTCTTACTGTGATGCAATCCGTAAGATGCTTTCTGTGACTGTGAGTACTCAAC 3101
QY 3236 AAGTCATTCTGAGAAATAGTGTATGCGGGAACCGAGTTGCTTGGCCCGCTCAATAGG 3295
Db 3102 AAGTCATTCTGAGAAATAGTGTATGCGGGAACCGAGTTGCTTGGCCCGCTCAATAGG 3161
QY 3296 GATTAATCCGCGCACATAGCAGAACTTTAAAGTCTCATCATTTGAAAAAGTTCTTCG 3355
Db 3162 GATTAATAGTATGACATAGCAGAACTTTAAAGTCTCATCATTTGAAAAAGTTCTTCG 3221
QY 3356 GGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCACTCGT 3415
Db 3222 GGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCACTCGT 3281
QY 3416 GCACCCAACTGATCTTCAGCATCTTTTACTTTTCAACCAAGCTTCTGGGTGACAAAAACA 3475
Db 3282 GCACCCAACTGATCTTCAGCATCTTTTACTTTTCAACCAAGCTTCTGGGTGACAAAAACA 3341
QY 3476 GGAAGGCAAAATGCGCAAAATGAGGAATAGGGCGACACGGAATGTTGAATCTCAT 3535
Db 3342 GGAAGGCAAAATGCGCGCAAAATGAGGAATAGGGCGACACGGAATGTTGAATCTCAT 3401
QY 3536 CTCTTCTTTTCAATATATATATGAAGCATTTATCAGGTTATGTTCTCATGAGCGATAC 3595
Db 3402 CTCTTCTTTTCAATATATATATGAAGCATTTATCAGGTTATGTTCTCATGAGCGATAC 3461
QY 3596 ATATTGAAATGATTTAGAAAAATTAACAATAGGGGTTCCGCGACATTTCCCGAAAA 3655
Db 3462 ATATTGAAATGATTTAGAAAAATTAACAATAGGGGTTCCGCGACATTTCCCGAAAA 3521
QY 3656 GTGCCAC 3662
Db 3522 GTGCCAC 3528
```

Search completed: January 26, 2005, 03:15:10  
Job time : 1771.03 secs

**This Page Blank (uspto)**



```
Db      1 CTAATGTGTAAGGCTTAATATTTGTTAAATTCGGTTAAATTTTGTAAATCAGCTC 60
Qy      61 ATTTTAAACCATATAGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAAATGACCGA 120
Db      61 ATTTTAAACCATATAGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAAATGACCGA 120
Qy      121 GATAGGGTTGAGTGTGTTCCAGTTTGGACAAGAGTCCACTATTTAAAGAACTGTGATC 180
Db      121 GATAGGGTTGAGTGTGTTCCAGTTTGGACAAGAGTCCACTATTTAAAGAACTGTGATC 180
Qy      181 CAACGTCAAAAGGCGAAAAACCGTCTATCAGGGCGATGCGCCACTAGGTAAACATACACC 240
Db      181 CAACGTCAAAAGGCGAAAAACCGTCTATCAGGGCGATGCGCCACTAGGTAAACATACACC 240
Qy      241 CTAAATCAAGTTTGTGGGGTGCAGAGTGCCTTAAAGCACTAAATCGAAACCTTAAAGGAG 300
Db      241 CTAAATCAAGTTTGTGGGGTGCAGAGTGCCTTAAAGCACTAAATCGAAACCTTAAAGGAG 300
Qy      301 CCCCCGATTTTGAAGCTTGAACGGGAAAGCCGGGAAAGCTGGCGAAGAAAGGAGAA 360
Db      301 CCCCCGATTTTGAAGCTTGAACGGGAAAGCCGGGAAAGCTGGCGAAGAAAGGAGAA 360
Qy      361 AGCGAAAGAGCGGGCGCTAGGGGCGTGGCAAGTGAAGCGTGCACGCTCGCGTAAACAC 420
Db      361 AGCGAAAGAGCGGGCGCTAGGGGCGTGGCAAGTGAAGCGTGCACGCTCGCGTAAACAC 420
Qy      421 CACACCCGCGCGCTTAATGCGCGCTACAGGGCGGCTCCATTGGCCATTGAGGCTGCG 480
Db      421 CACACCCGCGCGCTTAATGCGCGCTACAGGGCGGCTCCATTGGCCATTGAGGCTGCG 480
Qy      481 CAACGTGTGGGAAGGGGAGATCGGTGCGGCGCTTTCGCTATTAAGCCAGCTGGGAAAG 540
Db      481 CAACGTGTGGGAAGGGGAGATCGGTGCGGCGCTTTCGCTATTAAGCCAGCTGGGAAAG 540
Qy      541 GGGATGTCTGCAAGGCGATTAAGTGGGTAAACGCGAGGGTTTCCGCTACAGCACTTG 600
Db      541 GGGATGTCTGCAAGGCGATTAAGTGGGTAAACGCGAGGGTTTCCGCTACAGCACTTG 600
Qy      601 TAAACGAGCGGCGAGGAGCGGCGCTGTTCACTTCACTTTTGAACCCGTGAGAGAG 660
Db      601 TAAACGAGCGGCGAGGAGCGGCGCTGTTCACTTCACTTTTGAACCCGTGAGAGAG 660
Qy      661 GGAAGCTCGCGGTGCAAAATGTGTTTACAGGTAATGAGAGTAAGTAAGTGTCCAGAC 720
Db      661 GGAAGCTCGCGGTGCAAAATGTGTTTACAGGTAATGAGAGTAAGTAAGTGTCCAGAC 720
Qy      721 GCTGAGAAACGAGCTAGATTAACCTAGAAAGATTAATCATATGTGACGTAA 780
Db      721 GCTGAGAAACGAGCTAGATTAACCTAGAAAGATTAATCATATGTGACGTAA 780
Qy      781 AGATTAATATGCGTAAATTTGACCATGAGATCTGTATACGACTCATATAGGGCGAAT 840
Db      781 AGATTAATATGCGTAAATTTGACCATGAGATCTGTATACGACTCATATAGGGCGAAT 840
Qy      841 TGGGTACCGGGCCCCCTCGAGGTGCAAGGTATCAATTAAGCTTGAATCGAATCTGCG 900
Db      841 TGGGTACCGGGCCCCCTCGAGGTGCAAGGTATCAATTAAGCTTGAATCGAATCTGCG 900
Qy      901 AGCCCCGGGGATCCACTAGTGTGTAGAGCGGCGCAACGCGGTGAGAGCTCCAGCTTTGT 960
Db      901 AGCCCCGGGGATCCACTAGTGTGTGTAGAGCGGCGCAACGCGGTGAGAGCTCCAGCTTTGT 960
Qy      961 TCCCTTATGAGAGGTTAATTGAATCCATGCGTCAATTTTACGAGACTATCTTTGAG 1020
Db      961 TCCCTTATGAGAGGTTAATTGAATCCATGCGTCAATTTTACGAGACTATCTTTGAG 1020
Qy      1021 GGTAAATCTAGTGCATCAGATCAATATGTCGGGTCTTTTTCGGGTGAGTATGCC 1080
Db      1021 GGTAAATCTAGTGCATCAGATCAATATGTCGGGTCTTTTTCGGGTGAGTATGCC 1080
Qy      1081 CAAGCTGCGGCTATCTGGGCAATCGGGAGAAAGCCCGTGTCTTTTCCGCGAGGTTG 1140
Db      1081 CAAGCTGCGGCTATCTGGGCAATCGGGAGAAAGCCCGTGTCTTTTCCGCGAGGTTG 1140

Qy      1141 AAGCGCATGGAAGAGATTTCGAGAGATGACTGCTGTCATATGACGTTGAGCGAAAC 1200
Db      1141 AAGCGCATGGAAGAGATTTCGAGAGATGACTGCTGTCATATGAGCGTTGAGCGAAAC 1200
Qy      1201 GCAAGTTTAAACATGATGATTCGGAAGGTGTGGCAATGCAACGCTTTAAACGTTGA 1260
Db      1201 GCAAGTTTAAACATGATGATTCGGAAGGTGTGGCAATGCAACGCTTTAAACGTTGA 1260
Qy      1261 TCGTTACGGGCACTGGGGAATACAGATTTCGTGCGGCTTTTCCGGAACAAGTTCCGAT 1320
Db      1261 TCGTTACGGGCACTGGGGAATACAGATTTCGTGCGGCTTTTCCGGAACAAGTTCCGAT 1320
Qy      1321 TCAAGCCGGAAGCGCATCAGCAACCCGAACATATCCGGGAGACAGCCGGAATCTG 1380
Db      1321 TCAAGCCGGAAGCGCATCAGCAACCCGAACATATCCGGGAGACAGCCGGAATCTG 1380
Qy      1381 GTGAGCATTAATGACAGCGGTGCGGCGCTGGATATTAAGTACAGCAGAGACGGATAC 1440
Db      1381 GTGAGCATTAATGACAGCGGTGCGGCGCTGGATATTAAGTACAGCAGAGACGGATAC 1440
Qy      1441 CTGGCTGATGCGCGAGAAATGACATGATATCCCGGAGTTACCCGCGGCGGCTT 1500
Db      1441 CTGGCTGATGCGCGAGAAATGACATGATATCCCGGAGTTACCCGCGGCGGCTT 1500
Qy      1501 GGGCTAATGATGATGATGATGCTGTTTCTGTGTGAATTTGTATCGCTCAAAATTCACA 1560
Db      1501 GGGCTAATGATGATGATGATGCTGTTTCTGTGTGAATTTGTATCGCTCAAAATTCACA 1560
Qy      1561 CAACATACAGCGCGAAGCATAAAGTAAAGCCGTGGGGTCCCTAATGATGATGATGAT 1620
Db      1561 CAACATACAGCGCGAAGCATAAAGTAAAGCCGTGGGGTCCCTAATGATGATGATGAT 1620
Qy      1621 CACATTAATGCTGTGCTGCTACCTGCTTTCAGTGTGGGAAACCTGTGCTGACGCT 1680
Db      1621 CACATTAATGCTGTGCTGCTACCTGCTTTCAGTGTGGGAAACCTGTGCTGACGCT 1680
Qy      1681 GCAATTAATGATTCGGCCACAGCGCGGAGAGAGCGGTTTGGCTAATGATGATGATGAT 1740
Db      1681 GCAATTAATGATTCGGCCACAGCGCGGAGAGAGCGGTTTGGCTAATGATGATGATGAT 1740
Qy      1741 TTCTGCTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1800
Db      1741 TTCTGCTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1800
Qy      1801 CTCAAGCGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1860
Db      1801 CTCAAGCGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1860
Qy      1861 AGCAAAAGGCGAGCAAAAGCGAGAAACCGTAAAGGCGCGGTGATGATGATGATGATGAT 1920
Db      1861 AGCAAAAGGCGAGCAAAAGCGAGAAACCGTAAAGGCGCGGTGATGATGATGATGATGAT 1920
Qy      1921 TAGGCTCGGCGGCTGAGAGCATCAAAATTCGAGTCAAGTCAAGTCAAGTCAAGTCAAG 1980
Db      1921 TAGGCTCGGCGGCTGAGAGCATCAAAATTCGAGTCAAGTCAAGTCAAGTCAAGTCAAG 1980
Qy      1981 CCGGACAGATTAATTAAGATACAGAGGTTTCCCTGGAAGCTCCCTGATGATGATGATGAT 2040
Db      1981 CCGGACAGATTAATTAAGATACAGAGGTTTCCCTGGAAGCTCCCTGATGATGATGATGAT 2040
Qy      2041 TGTTCGAGCCCTGCGCTTACCGGATACGATGATGATGATGATGATGATGATGATGATGAT 2100
Db      2041 TGTTCGAGCCCTGCGCTTACCGGATACGATGATGATGATGATGATGATGATGATGATGAT 2100
Qy      2101 GCTTTCTATAGCTCAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2160
Db      2101 GCTTTCTATAGCTCAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2160
Qy      2161 GGGCTGTGTGACGAACCCCGGTTACAGCCGAGCCGCTTATCGGTAATCATATCG 2220
Db      2161 GGGCTGTGTGACGAACCCCGGTTACAGCCGAGCCGCTTATCGGTAATCATATCG 2220
```



QY 2221 TCTTAGTCCAAACCGGTAAAGACAGACTTAATCCGCACTGGAGAGCACTGGTAAG 2280  
 DB 2221 TCTTAGTCCAAACCGGTAAAGACAGACTTAATCCGCACTGGAGAGCACTGGTAAG 2280  
 QY 2281 GATTAGCAGACGAGGTATGTAGGCGGTCTACAGAGTCTTGAAGTGGGCTTAATA 2340  
 DB 2281 GATTAGCAGACGAGGTATGTAGGCGGTCTACAGAGTCTTGAAGTGGGCTTAATA 2340  
 QY 2341 CGGCTACCTAGAAAGACAGATTTGGTATCTGCGCTCTGCTGAAGCACTTACCTTGG 2400  
 DB 2341 CGGCTACCTAGAAAGACAGATTTGGTATCTGCGCTCTGCTGAAGCACTTACCTTGG 2400  
 QY 2401 AAAAAGATTGTAGCTCTTGTATCCGCAACAAACACCGGTGTAGCGGATTTT 2460  
 DB 2401 AAAAAGATTGTAGCTCTTGTATCCGCAACAAACACCGGTGTAGCGGATTTT 2460  
 QY 2461 TGTTCAGAGCAGAGATTACGCGAGAAAAAAGATCTCAGAAAGATCTTGTATCTT 2520  
 DB 2461 TGTTCAGAGCAGAGATTACGCGAGAAAAAAGATCTCAGAAAGATCTTGTATCTT 2520  
 QY 2521 TTCTACGGGGTCTGACGCTCAGTGAAGAAAACTCAGCTTAAGGATTTTGTATGAG 2580  
 DB 2521 TTCTACGGGGTCTGACGCTCAGTGAAGAAAACTCAGCTTAAGGATTTTGTATGAG 2580  
 QY 2581 ATTATCAAAAAGATCTTACCTAGATCTTTAAATTTAAATGAATTTTAAATCAAT 2640  
 DB 2581 ATTATCAAAAAGATCTTACCTAGATCTTTAAATTTAAATGAATTTTAAATCAAT 2640  
 QY 2641 CTAAAGTATATATAGTAACTTGTCTGACAGTTACCAATGCTTAATCAGTGAAGCAC 2700  
 DB 2641 CTAAAGTATATATAGTAACTTGTCTGACAGTTACCAATGCTTAATCAGTGAAGCAC 2700  
 QY 2701 TATCTCAGCACTCTGTCTATTTCTTCTTCAATAGTTCCTGACTCCCGCTGTATGAT 2760  
 DB 2701 TATCTCAGCACTCTGTCTATTTCTTCTTCAATAGTTCCTGACTCCCGCTGTATGAT 2760  
 QY 2761 AACTACGATACGGAGGGCTTACCACTGAGCCGAGCTGCAATGATACCGGAGACC 2820  
 DB 2761 AACTACGATACGGAGGGCTTACCACTGAGCCGAGCTGCAATGATACCGGAGACC 2820  
 QY 2821 ACGCTCAACGGCTCAGATTTATCAGCAATTAACAGCAGCAGCGAAGGCGAGCGAG 2880  
 DB 2821 ACGCTCAACGGCTCAGATTTATCAGCAATTAACAGCAGCAGCGAAGGCGAGCGAG 2880  
 QY 2881 AAGTGTCTCTGCACTTATCCGCTCCATCCAGTCTAATATTGTTCCGGAGACTAG 2940  
 DB 2881 AAGTGTCTCTGCACTTATCCGCTCCATCCAGTCTAATATTGTTCCGGAGACTAG 2940  
 QY 2941 AGTAAGTATGGCCAGTTAATAGTTTGGCGAAGCTTGGCAATGCTAAGAGCATCT 3000  
 DB 2941 AGTAAGTATGGCCAGTTAATAGTTTGGCGAAGCTTGGCAATGCTAAGAGCATCT 3000  
 QY 3001 GGTGTCAACGCTCTGTCTGTATGCTTCAATTCAGCTCCGCTTCCCAACATCAAGCG 3060  
 DB 3001 GGTGTCAACGCTCTGTCTGTATGCTTCAATTCAGCTCCGCTTCCCAACATCAAGCG 3060  
 QY 3061 AGTTACATGATCCCATGTTGTGCAAAAAAGCGGTATAGTCTCTTCCGATCTGT 3120  
 DB 3061 AGTTACATGATCCCATGTTGTGCAAAAAAGCGGTATAGTCTCTTCCGATCTGT 3120  
 QY 3121 TGTGAGAAGTATGGCCGAGTGTATCACTAGTGTATAGCAGACATGCAATATTC 3180  
 DB 3121 TGTGAGAAGTATGGCCGAGTGTATCACTAGTGTATAGCAGACATGCAATATTC 3180  
 QY 3181 TCTTACTGTCAATGCTCCGTAAGATCTTCTGTGATGCTGTAGTACTCAACCAAGTC 3240  
 DB 3181 TCTTACTGTCAATGCTCCGTAAGATCTTCTGTGATGCTGTAGTACTCAACCAAGTC 3240  
 QY 3241 ATTCTGAGAAATAGTGTATGCGCGACGAGATTGCTTTGCCGCGCTCAATACGGATAA 3300  
 DB 3241 ATTCTGAGAAATAGTGTATGCGCGACGAGATTGCTTTGCCGCGCTCAATACGGATAA 3300  
 QY 3301 TACCGCGCAATAGCAAACTTTAAAGTGTCTATCAATTTGGAAGAGTTCTTGGGGCG 3360

DB 3301 TACCGCGCAATAGCAAACTTTAAAGTGTCTATCAATTTGGAAGAGTTCTTGGGGCG 3360  
 QY 3361 AAACTCTCAAGATCTTACCGCTGTGAGATCCAGTTCGATGTATACCACTCTGTGACC 3420  
 DB 3361 AAACTCTCAAGATCTTACCGCTGTGAGATCCAGTTCGATGTATACCACTCTGTGACC 3420  
 QY 3421 CAATGATCTTCAAGATCTTCTTACTTTCACAGAGTTTCTGGGTGAGCAAAAAAGAGAG 3480  
 DB 3421 CAATGATCTTCAAGATCTTCTTACTTTCACAGAGTTTCTGGGTGAGCAAAAAAGAGAG 3480  
 QY 3481 GCAAAATGCGCAAAAAAGGAAATAGGCGACAGGAAATGTGAATTAATCAATCTCTT 3540  
 DB 3481 GCAAAATGCGCAAAAAAGGAAATAGGCGACAGGAAATGTGAATTAATCAATCTCTT 3540  
 QY 3541 CTTTTCAATATTTATGAGCAATTTACAGGTTATGTCTCATGAGCGGATACATATT 3600  
 DB 3541 CTTTTCAATATTTATGAGCAATTTATCAGGTTATGTCTCATGAGCGGATACATATT 3600  
 QY 3601 TGAATGTATTTAGAAAAATTAACAATAGGGGTTCCGCGCACTTTCCCGAAAAAGTCC 3660  
 DB 3601 TGAATGTATTTAGAAAAATTAACAATAGGGGTTCCGCGCACTTTCCCGAAAAAGTCC 3660  
 QY 3661 AC 3662  
 DB 3661 AC 3662

RESULT 2  
 US-10-211-079-18/c  
 ; Sequence 18, Application US/10211079  
 ; Publication No. US20030148315A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Padgett, Hal S.  
 ; APPLICANT: Vaeehongs, Andrew A.  
 ; APPLICANT: Voidani, Fakhrieh S.  
 ; APPLICANT: Smith, Mark L.  
 ; TITLE OF INVENTION: Nucleic Acid Molecules Encoding CEL I  
 ; TITLE OF INVENTION: Endonuclease and Methods of Use Thereof  
 ; FILE REFERENCE: P-LG 5381  
 ; CURRENT APPLICATION NUMBER: US/10/211,079  
 ; CURRENT FILING DATE: 2002-08-01  
 ; PRIOR APPLICATION NUMBER: US 10/098,155  
 ; PRIOR FILING DATE: 2002-03-14  
 ; NUMBER OF SEQ ID NOS: 29  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 18  
 ; LENGTH: 3637  
 ; TYPE: DNA  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: synthetic construct  
 US-10-211-079-18

Query Match 72.8%; Score 2665.4; DB 15; Length 3637;  
 Best Local Similarity 83.7%; Pred. No. 0;  
 Matches 3065; Conservative 0; Mismatches 571; Indels 27; Gaps 3;

QY 1 CTAAATGTAAAGCGTTAATATTTGTAAATTCGAGTTAAATTTTGTAAATCAGCTC 60  
 DB 3637 CTAAATGTAAAGCGTTAATATTTGTAAATTCGAGTTAAATTTTGTAAATCAGCTC 3578  
 QY 61 ATTTTAAACCAATAGGCGCAAAATCGGCAAAATCCCTTAATAATCAAAAGATAGACGA 120  
 DB 3577 ATTTTAAACCAATAGGCGCAAAATCGGCAAAATCCCTTAATAATCAAAAGATAGACGA 3518  
 QY 121 GATAGGTTAGTGTCTTGTCCATTGGAAACAAGTCCACTATTAAGAAAGTGAAGTCTC 180  
 DB 3517 GATAGGTTAGTGTCTTGTCTTGTGGAAACAAGTCCACTATTAAGAAAGTGAAGTCTC 3458  
 QY 181 CAACGTCAAGGCGCAAAACCGTCTATCAGGCGATGCGCCACTAGTGAACATCACC 240  
 DB 3457 CAACGTCAAGGCGCAAAACCGTCTATCAGGCGATGCGCCACTAGTGAACATCACC 3398

QY 241 CTAATCAAGTTTTTTGGGGTCAGAGTCCCTAAAGCACTAAATCCGAAACCCCTAANGAG 300  
Db 3397 CTAATCAAGTTTTTTGGGGTCAGAGTCCCTAAAGCACTAAATCCGAAACCCCTAANGAG 3338  
QY 301 CCCCCGATTTAGACTTGAACGGGGAAAGCCGGCGAACTGTGGGAAAGAAAGAGAA 360  
Db 3337 CCCCCGATTTAGACTTGAACGGGGAAAGCCGGCGAACTGTGGGAAAGAAAGAGAA 3278  
QY 361 AGGAAAGAGAGCGGGCGCTTAGGGCGCTGGCAAGTGTAGCGGTCAAGCTGGCGCTAACAC 420  
Db 3277 AGGAAAGAGAGCGGGCGCTTAGGGCGCTGGCAAGTGTAGCGGTCAAGCTGGCGCTAACAC 3218  
QY 421 CACACCCGCGCGCTTAATGCGCGCTACAGGGCGCGCTCCATTGCGCAATTCAGGCTGCG 480  
Db 3217 CACACCCGCGCGCTTAATGCGCGCTACAGGGCGCGCTCCATTGCGCAATTCAGGCTGCG 3158  
QY 481 CAACTGTTGGAGAGGGCGCATCGGTGCGGGCTCTTGCCTATTAGCGCAGCTGGCGAAAG 540  
Db 3157 CAACTGTTGGAGAGGGCGCATCGGTGCGGGCTCTTGCCTATTAGCGCAGCTGGCGAAAG 3098  
QY 541 GGGATGTCCTCAAGGCGATTAAGTGGGTAACGCGAGGTTTTTCCAGTCAAGAGCTTG 600  
Db 3097 GGGATGTCCTCAAGGCGATTAAGTGGGTAACGCGAGGTTTTTCCAGTCAAGAGCTTG 3038  
QY 601 TAAAAAGAGCGGCGATGAGCGCGCTCTGTTCAATTCAGTTTTGAAACCGGTGAGAGAG 660  
Db 3037 TAAAAAGAGCGGCGATGAGCGCGCTCTGTTCAATTCAGTTTTGAAACCGGTGAGAGAG 2978  
QY 661 GCAAGCTCGCGGTGCAAAATGTGTTTTACAGCGTGAATGAGCAAGATGCTGACAC 720  
Db 2977 CGCGGTGCGCGCGCTCTGTTCAATTCAGCGTGAATGAGCAAGATGCTGACAC 720  
QY 721 GCTGCAAGAACGCGAGCTAGATTAACTTAAGAAAGATATATGAGAGTGAAGTGA 780  
Db 2917 ATAGTTCATTCAGTCCATGTGTAATCCAGAGCACTTAACAATCAAGAGAGCAATGT 2858  
QY 781 AGATTAATC-ATGCGTAAATTAAGACGATGAGATCTGTAATACATCACTATAGGGCGAA 839  
Db 2857 GGTCAAGCTTTTGTGTGGGATCTTTGAAAGGCGAATGTGTGCAAGGTAATGTGTTG 2798  
QY 840 TTGGGTACCGGGCCCCCTCGAGGTCAAGGTATCGATTAAGCTTGAATTCGAATTCCTG 899  
Db 2797 CTGGTAAAGAGACAGGGCATGCGCAATTTGAGTAATTTGTGATATGCTGTGCTAAT 2738  
QY 900 CAGCCCGGGGATCACTAGTTCTAGACGCGCGCCACCGGGGTGAGCTCCAGCTTTG 959  
Db 2737 GAGCGATTCATCTTCAATGTGTGGGAAATTTGAAGTTAGCTTGAATTCATTTCTTT 2678  
QY 960 TTCCCTTATGAGGGTAAATTAATCCCAATGCGTCAATTTAAGGAGACTAATCTTTCTA 1019  
Db 2677 GTTGTCTGCGGTATGTATCATTTGTGTGAAGTTAAGTTAATCTGAGTTGTGTCCGA 2618  
QY 1020 GGGTTAATCTAGCTGCAATCAGAGATCATATGTCGGGCTTTTTTCCGGCTCAGTATGC 1079  
Db 2617 GAAATTTTCAATCTTTTAAATTAACCTTTAATCTGATAC-----GATTAAACAG 2563  
QY 1080 CCAAGCTGGGCTATCTGGGATCTGGGAGAGAAAGCCGTCCTTTTCCCGGAGGTT 1139  
Db 2562 GGTATCACTTTCAACCTTGAATCTTCAAGCAGCGCTTGTAGTTCCCGCATCTTTGAAGA 2503  
QY 1140 GAAGCGGATGAGAAAGTTTGGCGAGATGACTGCTGCTGATTTGAAGTGAAGCGAA 1199  
Db 2502 TATAGTCCGTTCTGTATCAATCTTGGGCAATGCACTTTGAAAGATGATCCGTTT 2443  
QY 1200 CGCAGCTTAACTATGATGATTCGGGAAGGTGGCATGCAAGCCTTTAAACGTTGAATG 1259  
Db 2442 CATATGATCCGATTAACGAGAAACATTTGAACACCATGAGAAATAGTGAACAAGTGT 2383  
QY 1260 TTGCTTACAGCACTCGGATACAGATTGCTGCGGGCTTTTCCGAGACAGTTCCGGAATG 2319  
Db 2382 TGGCATTGAAAGAGTATGTTTCCAGTGTGCAAAATTAATTAAGTTAAGTTAGTTCCGTA 2323

QY 1320 GTACGCCGAAACCGCATGACAAACCCGAACAATACCGGCGACAGCGGAACTGCCGTGC 1379  
Db 2322 TGTACATACCTTCAACCTCTCCACTGACAGAAATTTGTGCCATTAACTACATAC 2263  
QY 1380 GGTGTGCAATTAATGACAGCGGTGGCGCTGGATATTATGCTCAAGCAAGACGGTAT 1439  
Db 2262 TAATTCACAGAAAT-----TGGAGCAACTCCAGTGAAGAAAGTTCC 2224  
QY 1440 CCGGCTGGATAGCCGCGAAATAGACATGATATCCCGTGAATACCCGCGGGCGCT 1499  
Db 2223 TTCTCTTTACTACATGCTGATCCAGCTTTGTCTCTTTAAGAGGTTAATTTGGCGCT 2164  
QY 1500 TGGGTAATCATGATGATAGCTGTTTCTGTGTGAATTTGTTATCTGCTCAAAATTCAC 1559  
Db 2163 TGGGTAATCATGATGATAGCTGTTTCTGTGTGAATTTGTTATCTGCTCAAAATTCAC 2104  
QY 1560 ACAACATACAGCCGGAACGATTAAGTGAAGCTGGGGTGGCTTAATGATGAGCTTAC 1619  
Db 2103 ACAACATACAGCCGGAACGATTAAGTGAAGCTGGGGTGGCTTAATGATGAGCTTAC 2044  
QY 1620 TCACTTAATTTGGTGTGCTCACTGCGCGCTTTCAGTGGAAACCTGTGTGCGAGC 1679  
Db 2043 TCACTTAATTTGGTGTGCTCACTGCGCGCTTTCAGTGGAAACCTGTGTGCGAGC 1984  
QY 1680 TGCATTAATGAATCGGCGCAACCGCGGGAGAGCGGTTGCGTATTTGGCGCTTCCG 1739  
Db 1983 TGCATTAATGAATCGGCGCAACCGCGGGAGAGCGGTTGCGTATTTGGCGCTTCCG 1924  
QY 1740 CTTCCTCGCTCACTGACTCGCTGCGCTCGGCTGTTGGCGTGGCGGAGGAGTTCAGCTC 1799  
Db 1923 CTTCCTCGCTCACTGACTCGCTGCGCTCGGCTGTTGGCGTGGCGGAGGAGTTCAGCTC 1864  
QY 1800 ACTCAAGCGGTAATCGGTTATCCACAGATTCAGGGGATTAACGAGAGAAACATGT 1859  
Db 1863 ACTCAAGCGGTAATCGGTTATCCACAGATTCAGGGGATTAACGAGAGAAACATGT 1804  
QY 1860 GAGCAAAAGGCGAGCAAAAGCCAGAACCTGTAAGAGCGCGTGTGCTGCGCTTTTCC 1919  
Db 1803 GAGCAAAAGGCGAGCAAAAGCCAGAACCTGTAAGAGCGCGTGTGCTGCGCTTTTCC 1744  
QY 1920 ATAGGCTCGCGCCCGCGTACAGACATCAAAAATCGACGCTCAAGTCAAGAGTGGGAA 1979  
Db 1743 ATAGGCTCGCGCCCGCGTACAGACATCAAAAATCGACGCTCAAGTCAAGAGTGGGAA 1864  
QY 1980 ACCGCAAGAGCTAATAAGATACAGGCGTTTCCCGTGAAGCTCCCTGTGCGCTC 2039  
Db 1683 ACCGCAAGAGCTAATAAGATACAGGCGTTTCCCGTGAAGCTCCCTGTGCGCTC 1624  
QY 2040 CTGTTCCGACCCCTGCGCGCTTACCGGATCTGTGCTGCGCTTTCTCCCTCGGAAAGCGTG 2099  
Db 1623 CTGTTCCGACCCCTGCGCGCTTACCGGATCTGTGCTGCGCTTTCTCCCTCGGAAAGCGTG 1564  
QY 2100 CGCTTTCTATAGCTCAAGCTGTAGATCTCAAGTTCCGATGAGTGTGCTGCTCAAC 2159  
Db 1563 CGCTTTCTATAGCTCAAGCTGTAGATCTCAAGTTCCGATGAGTGTGCTGCTCAAC 1504  
QY 2160 TGGGCTGTGTCAGAAACCCCGCTTACCGGCGACCGCTGCGCTTATCCGTAATC 2219  
Db 1503 TGGGCTGTGTCAGAAACCCCGCTTACCGGCGACCGCTGCGCTTATCCGTAATC 1444  
QY 2220 GTCTTGAATCAACCGCGTAAGACAGACTTAATCGCACTGGCAGAGCACTGGTAACA 2279  
Db 1443 GTCTTGAATCAACCGCGTAAGACAGACTTAATCGCACTGGCAGAGCACTGGTAACA 1384  
QY 2280 GGATTAACAGAGGATATGAGGGGCTTACAGAGTTCTTGAAGTGTGCGCTAATC 2339  
Db 1383 GGATTAACAGAGGATATGAGGGGCTTACAGAGTTCTTGAAGTGTGCGCTAATC 1324  
QY 2340 ACGGCTACACTAAGAGACAGATTTGTATCTGCGCTCTGCTGAAGCACTTACCTTCG 2399  
Db 1323 ACGGCTACACTAAGAGACAGATTTGTATCTGCGCTCTGCTGAAGCACTTACCTTCG 1264  
QY 2400 GAAAAAGATTGTAGCTTGTATCCGGCAACAAACACCGCTGTAGCGGTGTTT 2459

Db 1263 GAAAAGGTTGTAAGCTCTTGATCCGGCAAAACACCGCTGTGACGGTGGTTTTT 1204  
Qy 2460 TTGTTTGAACGACGAGATTAACGGCGAGAAAAAAGATCTCAAGAGATCTTTGATCT 2519  
Db 1203 TTGTTTGAACGACGAGATTAACGGCGAGAAAAAAGATCTCAAGAGATCTTTGATCT 1144  
Qy 2520 TTTTCAAGGGGCTGACGCTCAGTGGAACGAAAACTCAGTTAAGGATTTTGGTCAAGA 2579  
Db 1143 TTTTCAAGGGGCTGACGCTCAGTGGAACGAAAACTCAGTTAAGGATTTTGGTCAAGA 1084  
Qy 2580 GATTATCAAAAAGGATCTTCAACCTTAGATCTTTTAAATTAATAAGTATTTAAATCAA 2639  
Db 1083 GATTATCAAAAAGGATCTTCAACCTTAGATCTTTTAAATTAATAAGTATTTAAATCAA 1024  
Qy 2640 TCTAAGTATATATAGATTAACCTTGCTGACAGTTACCAATGCTTAATCAGTAGGAC 2699  
Db 1023 TCTAAGTATATATAGATTAACCTTGCTGACAGTTACCAATGCTTAATCAGTAGGAC 964  
Qy 2700 CTATCTCAGCGATCTGTCTATTTTGTTCATCATATGTTGCTGACTCCCGCTGCTGA 2759  
Db 963 CTATCTCAGCGATCTGTCTATTTTGTTCATCATATGTTGCTGACTCCCGCTGCTGA 904  
Qy 2760 TAACTACGATACGGGAGGCTTACCATCTGGCCCAAGTGTGCAATGATACGGAGAC 2819  
Db 903 TAACTACGATACGGGAGGCTTACCATCTGGCCCAAGTGTGCAATGATACGGAGAC 844  
Qy 2820 CAGCTCACCGGCTTCAGATTTATCAGCAATTAACCAAGCCGGAAGGGCCGACGCA 2879  
Db 843 CAGCTCACCGGCTTCAGATTTATCAGCAATTAACCAAGCCGGAAGGGCCGACGCA 784  
Qy 2880 GAACTGTCTCTGCAACTTTATCCGCTCCATCTCATATTAATTTGTCGGGAAAGCTA 2939  
Db 783 GAACTGTCTCTGCAACTTTATCCGCTCCATCTCATATTAATTTGTCGGGAAAGCTA 724  
Qy 2940 GAGTAAGTATGTCGCGCAGTTAATAGTTGCGCAAGTGTGCGCATGCTCAAGGATCG 2999  
Db 723 GAGTAAGTATGTCGCGCAGTTAATAGTTGCGCAAGTGTGCGCATGCTCAAGGATCG 664  
Qy 3000 TGGTGTCAAGCTCTGTGTTGGTATAGCTTATCAAGCTCCGCTCCCAAGATCAAGGC 3059  
Db 663 TGGTGTCAAGCTCTGTGTTGGTATAGCTTATCAAGCTCCGCTCCCAAGATCAAGGC 604  
Qy 3060 GAGTTACATGATCCCCCATGTTTGTGCAAAAAAGCGGTAGCTCTTCCGCTCCGATCG 3119  
Db 603 GAGTTACATGATCCCCCATGTTTGTGCAAAAAAGCGGTAGCTCTTCCGCTCCGATCG 544  
Qy 3120 TTGTCAGAAATAGTTGGCCGAGTGTATCATCTCATAGGTTATGCGACACTGCATATT 3179  
Db 543 TTGTCAGAAATAGTTGGCCGAGTGTATCATCTCATAGGTTATGCGACACTGCATATT 484  
Qy 3180 CTCTTACTGTATGTCATCGTAAGATCTTTCTGTGATCTGTGATGACTCAACCAAGT 3239  
Db 483 CTCTTACTGTATGTCATCGTAAGATCTTTCTGTGATCTGTGATGACTCAACCAAGT 424  
Qy 3240 CATCTGTGAATAGTGTATGCGGCAACGAGTTGCTCTTGGCCGCGTCAATAGCGGATA 3299  
Db 423 CATCTGTGAATAGTGTATGCGGCAACGAGTTGCTCTTGGCCGCGTCAATAGCGGATA 364  
Qy 3300 ATACCGGCGCATATGACGAACTTTAAAGTGTATCATATGGAAGAGTCTTCCGGGGC 3359  
Db 363 ATACCGGCGCATATGACGAACTTTAAAGTGTATCATATGGAAGAGTCTTCCGGGGC 304  
Qy 3360 GAAAACTCTCAGGATCTTACCGCTGTGTAGATCCAGTTGATGATCAACCTCGTGAC 3419  
Db 303 GAAAACTCTCAGGATCTTACCGCTGTGTAGATCCAGTTGATGATCAACCTCGTGAC 244  
Qy 3420 CCAACTGATCTTCAAGCATCTTTTACTTCAACGAGTTTCTGGGTGAGCAAAAACAGGAA 3479  
Db 243 CCAACTGATCTTCAAGCATCTTTTACTTCAACGAGTTTCTGGGTGAGCAAAAACAGGAA 184  
Qy 3480 GGCAAAATGCGCAAAAAGGAAATPAGGCGACACGGAATGTTGAATACCTCAATACCT 3539

Db 183 GGCAAAATGCCGCAAAAAGGAAATPAGGCGACACGGAATGTTGAATCTCAATCT 124  
Qy 3540 TCTTTTTCATATATATTTGAAGCATTTTATCAGGGTTATTTGTCTATGACGGATACATAT 3599  
Db 123 TCTTTTTCATATATATTTGAAGCATTTTATCAGGGTTATTTGTCTATGACGGATACATAT 64  
Qy 3600 TTGAATGATTTAGAAAAATPAAACAAATPAGGGGTTCCGGCAGATTTCCCGAAGATGC 3659  
Db 63 TTGAATGATTTAGAAAAATPAAACAAATPAGGGGTTCCGGCAGATTTCCCGAAGATGC 4  
Qy 3660 CAC 3662  
Db 3 CAC 1  
  
RESULT 3  
US-10-356-708-32/c  
; Sequence 32, Application US/10356708  
; Publication No. US20030157682A1  
; GENERAL INFORMATION:  
; APPLICANT: Large Scale Biology Corporation  
; TITLE OF INVENTION: MISMATCH ENDONUCLEASES AND METHODS OF USE THEREOF  
; FILE REFERENCE: P-1G 10100  
; CURRENT APPLICATION NUMBER: US/10/356,708  
; PRIOR FILING DATE: 2003-01-31  
; PRIOR APPLICATION NUMBER: 60/353,722  
; PRIOR FILING DATE: 2000-02-01  
; PRIOR APPLICATION NUMBER: 10/098,155  
; PRIOR FILING DATE: 2002-03-14  
; PRIOR APPLICATION NUMBER: 10/211,079  
; PRIOR FILING DATE: 2002-08-01  
; NUMBER OF SEQ ID NOS: 35  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 32  
; LENGTH: 3637  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Encodes Cycle 3 BFP gene  
US-10-356-708-32  
  
Query Match 72.8%; Score 2665.4; DB 15; Length 3637;  
Best Local Similarity 83.7%; Pred. No. 0;  
Matches 3065; Conservative 0; Mismatches 511; Indels 27; Gaps 3;  
  
Qy 1 CTAATTTGTAAGCTTATATATTTTGTAAATTCGCTTAAATTTTGTAAATCAGCTC 60  
Db 3637 CTAATTTGTAAGCTTATATATTTTGTAAATTCGCTTAAATTTTGTAAATCAGCTC 3578  
Qy 61 ATTTTAAACCAATAGCGCGGAATCGGCAAAATCCCTTATTAATCAAAAGATGACCGA 120  
Db 3577 ATTTTAAACCAATAGCGCGGAATCGGCAAAATCCCTTATTAATCAAAAGATGACCGA 3518  
Qy 121 GATAGGTTAGAGTGTTCCTCAAGTTGGAAACAAGTCCACTAATTAAGAACGTGACCTC 180  
Db 3517 GATAGGTTAGAGTGTTCCTCAAGTTGGAAACAAGTCCACTAATTAAGAACGTGACCTC 3458  
Qy 181 CAACGTCAAGGGCGGAAAAACCGTCTATCAGGGCGATGCGCCACTACGTGAACCATGACC 240  
Db 3457 CAACGTCAAGGGCGGAAAAACCGTCTATCAGGGCGATGCGCCACTACGTGAACCATGACC 3398  
Qy 3397 CTAATCAAGATTTTGTGGGGTCAAGTCCCTTAAGAACAATAATCGAAGAGAGAG 3338  
Db 301 CCCCCGATTTAGAGCTTGAAGGGGGAAGCCGGGGAACGTGCGGGAAGGAAGGAGAA 360  
Qy 3337 CCCCCGATTTAGAGCTTGAAGGGGGAAGCCGGGGAACGTGCGGGAAGGAAGGAGAA 3278  
Db 361 AGCGAAAGAGCGGGCGCTAGGGCGCTGAGCAAGTGTAGCGGTCAACGTGCGGTAAACAC 420  
Qy 3277 AGCGAAAGAGCGGGCGCTAGGGCGCTGAGCAAGTGTAGCGGTCAACGTGCGGTAAACAC 3218

QY 421 CACACCCGCGCGCTTAATGCGCCGCTACAGGGCGCGCTCCCAATGCGCAATTCAGGCTCG 480  
DB |||||  
DB 3217 CACACCGCGCGCTTAATGCGCGCTACAGGGCGCGCTCCCAATGCGCAATTCAGGCTCG 3158  
QY 481 CAACTGTTGGGAAGGGGCGATCGGTGCGGGCTCTTCGCTATTACGCCGACTGGCGAAAG 540  
DB 3157 CAACTGTTGGGAAGGGGCGATCGGTGCGGGCTCTTCGCTATTACGCCGACTGGCGAAAG 3098  
QY 541 GGGATGTCGCAAGGCGATTAAAGTTGGGTAAACGCGAGGGTTTCCAGTACAGACGTTG 600  
DB 3097 GGGATGTCGCAAGGCGATTAAAGTTGGGTAAACGCGAGGGTTTCCAGTACAGACGTTG 3038  
QY 601 TAAAAAGACGGCCAGTAGAGCGCGCTTCGTTCAATGACGTTTGAACCCGTGAGAGCGG 660  
DB 3037 TAAAAAGACGGCCAGTAGAGCGCGCTTCGTTCAATGACGTTTGAACCCGTGAGAGCGG 2978  
QY 661 GCAGACTCGCGGTGCAAAATGTGTTTACAGCGTGAAGAGAGATGAAGATGCGCAGAC 720  
DB 2977 CCGCGGTGCGCGCGCTCTAGAACTAGTGTATCCCGCGGCTGAGGAAATCTTATTTGT 2918  
QY 721 GCTGCAAAACAGCAGCTAGATTAAACCTAGAAAGATATCATATTGTGACGTATTA 780  
DB 2917 ATAGTTCAATCCATGCGCATGTGTATCCAGACAGAGATTACAACTAAGAGACCATGT 2858  
QY 781 AGATATATC-ATGCGTAAATTTGACGCAATGGATCTGTAAATACGACTCACTATAGGGCGCA 839  
DB 2857 GGTACGCTTTTCGTTGGGATCTTTGAAAGGGCAGATTGTGTGACAGTAAATGTGT 2798  
QY 840 TTGGGTACGGGCGCGCGCTCGAGGTGACGATGATGATGATGATGATGATGATGATGAT 899  
DB 2797 CTGGTAAAGAGACAGGGCCATCGCAATTGGAGTATTTGTGATTAATGTCTGCTAGTT 2738  
QY 900 CAGCCCGGGGATTCACCTAGTTCTAGAGCGCGCGCCACCGCGTGAAGCTCCAGCTTTG 959  
DB 2737 GAACGATTCATCTTCAATGTTGTGGCAATTTTGAAGTAGCTTGAATTCATTTCTTT 2678  
QY 960 TTCCCTTAATGAGGTTAATTAATGATCCCATGCTCAATTTTACGACGATCTTTCTA 1019  
DB 2677 GTTTGTCGCGGTGATGATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 2618  
QY 1020 GGGTATATCTAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1079  
DB 2617 GAATGTTTCATCTTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1080  
QY 1080 CCAAGTGGCGCTATCTGGGACATCGGGGAGGAAAGAACCGGTGCTTTTCCCGAGGTT 1139  
DB 2562 GGTATCATCTTCAACTTGACCTTGACGCGGTCTTGAATCCCGTCACTTTTGAAGA 2503  
QY 1140 GAAGCGGATGGAAGAGTTGCGGAGATGACGCTGCTGATGATGATGATGATGATGATGAT 1199  
DB 2502 TATAGTGGTTCCTGTATACATCTTGCGGACATGGCACTCTGAAAGAGTCAAGCTT 2443  
QY 1200 CGCAGTTTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1259  
DB 2442 CATATGATCCGATTAAGAGAAAGCATGAAACCATAGAGAAAGTAGTAACAAGTGT 2383  
QY 1260 TTGCTTACAGGCACTGGAATACAGATTCGTCCGCGCTTTTCGGAACAAGTTCCGGAATG 1319  
DB 2382 TGGCATGGAACAGGTGTTTCAAGTAGTGAATTAATTAATTAATTAATTAATTAATTAAT 2323  
QY 1320 GTGAGCCGGAAGCGCATCAGCAACCGGAACAATAACGGGCAAGCGCGGAATCTGCGTGC 1379  
DB 2322 TGTAGCATCCTTCAACCTCTCTCACTGACAGAAATTTTGTCCATTAATCAATCAATC 2263  
QY 1380 GGTGTACAGATTAAATGACAGCGGTGCGGCTGGGATTAATGATGATGATGATGATGATGAT 1439  
DB 2262 TAATTAACAAGAT-----TGGGCAACTCAAGTGAAGAAAGTTTC 2224  
QY 1440 CCGGTGATGATGCGGAGAAATGACATGATACCCCGTGAAGTACCCGCGCGCGCT 1499  
DB 2223 TTCTCTTATCATCGTACCACTTTTGTTCCTTTATGAGGGTTAATTTGCGCGCT 2164  
QY 1500 TGGCGTATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1559

DB 2163 TGGCGTATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2104  
QY 1560 ACACATATCAGACCGGAAGCATTAAGTAAAGCTGGGTGCTTATGATGATGATGATGATGATGAT 1619  
DB 2103 ACACATATCAGACCGGAAGCATTAAGTAAAGCTGGGTGCTTATGATGATGATGATGATGATGAT 2044  
QY 1620 TCACATTAATGCGGTGCGCTCACTGCGCGCTTTCCAGTGGGAAACCTGTGTCGAGC 1679  
DB 2043 TCACATTAATGCGGTGCGCTCACTGCGCGCTTTCCAGTGGGAAACCTGTGTCGAGC 1984  
QY 1680 TGCATTAATGATCGGCAACGCGCGGAGAGGCGGTTTGGGTATGGGCGCTCTTCG 1739  
DB 1983 TGCATTAATGATCGGCAACGCGCGGAGAGGCGGTTTGGGTATGGGCGCTCTTCG 1924  
QY 1740 CTTCCTGCTCACTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1799  
DB 1923 CTTCCTGCTCACTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1864  
QY 1800 ACTCAAGCGGTAAATCGGTATATCAGAGATCAGAGATCAGAGATCAGAGATCAGAGATCAGAGAT 1859  
DB 1863 ACTCAAGCGGTAAATCGGTATATCAGAGATCAGAGATCAGAGATCAGAGATCAGAGATCAGAGAT 1804  
QY 1860 GAGCAAAAGCGGAGAAAGCGGAAACGTTAAAGCGCGCTTGTGCTGCTGCTGCTGCTGCTGCTGCT 1919  
DB 1803 GAGCAAAAGCGGAGAAAGCGGAAACGTTAAAGCGCGCTTGTGCTGCTGCTGCTGCTGCTGCTGCT 1744  
QY 1920 ATAGGCTCGCGCGCTGACGAGATCAAAAATGAGGCTCAATGATGATGATGATGATGATGATGAT 1979  
DB 1743 ATAGGCTCGCGCGCTGACGAGATCAAAAATGAGGCTCAATGATGATGATGATGATGATGATGAT 1684  
QY 1980 ACCCGACAGACTTAAGATATACAGGCGCTTTCCCTGGAAGCTTCTGTCGCTGCTGCTGCTGCTGCT 2039  
DB 1683 ACCCGACAGACTTAAGATATACAGGCGCTTTCCCTGGAAGCTTCTGTCGCTGCTGCTGCTGCTGCT 1624  
QY 2040 CTGTTTCGACCTGCGCTTACCGGATACCTGTCGCGCTTTCTCCCTTCCGGAAGCGTGG 2099  
DB 1623 CTGTTTCGACCTGCGCTTACCGGATACCTGTCGCGCTTTCTCCCTTCCGGAAGCGTGG 1564  
QY 2100 CGCTTTCATAGCTACAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2159  
DB 1563 CGCTTTCATAGCTACAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1504  
QY 2160 TGGGCTGTGTGACGAACCCCGCTTACGCGCGGACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2219  
DB 1503 TGGGCTGTGTGACGAACCCCGCTTACGCGCGGACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1444  
QY 2220 GTCTTGAATCAACCGGTAAGACGACTTATGCGCATCTGGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 2279  
DB 1443 GTCTTGAATCAACCGGTAAGACGACTTATGCGCATCTGGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1384  
QY 2280 GATTTAGCAGACGAGTATGAGCGGTGCTACAGAGTTCTTGAAGTGGTGGCTTAATCT 2339  
DB 1383 GATTTAGCAGACGAGTATGAGCGGTGCTACAGAGTTCTTGAAGTGGTGGCTTAATCT 1324  
QY 2340 ACGGCTACACTAGAAAGCAGTATTTGGTATCTGCGCTGCTGTAAGCCAGTTACTTCG 2399  
DB 1323 ACGGCTACACTAGAAAGCAGTATTTGGTATCTGCGCTGCTGTAAGCCAGTTACTTCG 1264  
QY 2400 GAAAAAGGTTGTAAGCTCTTGAATCCGGAACAACCCGCTGTAAGCCAGTTACTTCG 1204  
DB 1263 GAAAAAGGTTGTAAGCTCTTGAATCCGGAACAACCCGCTGTAAGCCAGTTACTTCG 2519  
QY 2460 TTGTTTCAAGCAGCAGATTAAGCGGAGAAAAAGATCTCAAGAAATCTTGTGATCT 1144  
DB 1203 TTGTTTCAAGCAGCAGATTAAGCGGAGAAAAAGATCTCAAGAAATCTTGTGATCT 2579  
QY 2520 TTCTTACGCGGTGTAAGCTCACTGATGAGAAAGAACTCAAGTTAAGGATTTTGTGATGTA 1084  
DB 1143 TTCTTACGCGGTGTAAGCTCACTGATGAGAAAGAACTCAAGTTAAGGATTTTGTGATGTA 2639  
QY 2580 GATTAATCAAAAGATCTTCACTGATCTTTTAATTAATTAATTAATTAATTAATTAATTAATCA 2639

Db 1083 GATTATCAAAAAGATCTTACACCTAGATCTTTAAATTAATAAGATTTAAATCA 1024  
Qy 2640 TCTAAAGATATATAGATTAACCTTGTCTGACAGTTACCAATGCTTAATCACTGAGC 2699  
Db 1023 TCTAAAGATATATAGATTAACCTTGTCTGACAGTTACCAATGCTTAATCACTGAGC 964  
Qy 2700 CTATCTGACGATCTGTCTATTTGCTTATCCATCCATGAGTGGCCGCTGAGTGA 2759  
Db 963 CTATCTGACGATCTGTCTATTTGCTTATCCATCCATGAGTGGCCGCTGAGTGA 904  
Qy 2760 TAACTACGATACGGAGGGCTTACCATCTGGCCCAAGCTGCAATGATACGGAGC 2819  
Db 903 TAACTACGATACGGAGGGCTTACCATCTGGCCCAAGCTGCAATGATACGGAGC 844  
Qy 2820 CACGCTCACCGGCTCCAGATTTATAGCAATTAACACCGACCGGAGGGCCGCA 2879  
Db 843 CACGCTCACCGGCTCCAGATTTATAGCAATTAACACCGACCGGAGGGCCGCA 784  
Qy 2880 GAAAGTGTCTGCAACTTTATCCGCTCCATCCAGTCTTAAATTTGTTGCCGGAGGCTA 2939  
Db 783 GAAAGTGTCTGCAACTTTATCCGCTCCATCCAGTCTTAAATTTGTTGCCGGAGGCTA 724  
Qy 2940 GAGTAAGTGTCTGCAACTTTATCCGCTCCATCCAGTCTTAAATTTGTTGCCAGGCTG 2999  
Db 723 GAGTAAGTGTCTGCAACTTTATCCGCTCCATCCAGTCTTAAATTTGTTGCCAGGCTG 664  
Qy 3000 TGTGTCTACGCTCGCTGTTGGTATGCTTCAATGCTCCGGTCCCAAGATCAAGC 3059  
Db 663 TGTGTCTACGCTCGCTGTTGGTATGCTTCAATGCTCCGGTCCCAAGATCAAGC 604  
Qy 3060 GAGTTACATGATCCCCCAATGTTGTCAGAAAAAGGGGTAGGCTCTCGGCTCTCCGATCG 3119  
Db 603 GAGTTACATGATCCCCCAATGTTGTCAGAAAAAGGGGTAGGCTCTCGGCTCTCCGATCG 544  
Qy 3120 TTGTTCAGAAATGATGTGGCCGAGTGTATCATCTCATGCTTATGTCAGACATGATAT 3179  
Db 543 TTGTTCAGAAATGATGTGGCCGAGTGTATCATCTCATGCTTATGTCAGACATGATAT 484  
Qy 3180 CTCTTACCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3239  
Db 483 CTCTTACCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 424  
Qy 3240 CATCTGAGAAATGATGTGGCCGAGTGTATCATCTCATGCTTATGTCAGACATGATAT 3299  
Db 423 CATCTGAGAAATGATGTGGCCGAGTGTATCATCTCATGCTTATGTCAGACATGATAT 364  
Qy 3300 ATACCGGCGCATATGAGCAAACTTTAAAGTGTCTGATGATGGAAGAGTCTTCCGAGC 3359  
Db 363 ATACCGGCGCATATGAGCAAACTTTAAAGTGTCTGATGATGGAAGAGTCTTCCGAGC 304  
Qy 3360 GAAAACTTCAAGGATCTTACCGCTGTGAGATCCAGTTCAGATGTAACCTCCTGAC 3419  
Db 303 GAAAACTTCAAGGATCTTACCGCTGTGAGATCCAGTTCAGATGTAACCTCCTGAC 244  
Qy 3420 CCAATGATCTTCAAGGATCTTACCGCTGTGAGATCCAGTTCAGATGTAACCTCCTGAC 3479  
Db 243 CCAATGATCTTCAAGGATCTTACCGCTGTGAGATCCAGTTCAGATGTAACCTCCTGAC 184  
Qy 3480 GGGAAATGCGGCAAAAAAGGAAATAGGGGCAACGGAATGTTGATATCTACTCT 3539  
Db 183 GGGAAATGCGGCAAAAAAGGAAATAGGGGCAACGGAATGTTGATATCTACTCT 124  
Qy 3540 TCTTTTTCATATTTATGAGCATTTATCAGGGTATTTGCTCATGAGCGGATCATAT 3599  
Db 123 TCTTTTTCATATTTATGAGCATTTATCAGGGTATTTGCTCATGAGCGGATCATAT 64  
Qy 3600 TTGAATGATTTAGAAAAATAAACAATAGGGGTTCCGGGCAATTTCCCGAAAAAGTC 3659  
Db 63 TTGAATGATTTAGAAAAATAAACAATAGGGGTTCCGGGCAATTTCCCGAAAAAGTC 4  
Qy 3660 CAC 3662  
Db 3 CAC 1

RESULT 4  
US-10-280-913A-32/c  
; Sequence 32, Application US/10280913A  
; Publication No. US20040110130A1  
; GENERAL INFORMATION:  
; APPLICANT: Large Scale Biology Corporation  
; TITLE OF INVENTION: POPULATION OF POLYNUCLEOTIDE SEQUENCE VARIANTS  
; FILE REFERENCE: LSBC-PLG5392-CIP  
; CURRENT APPLICATION NUMBER: US/10/280, 913A  
; PRIOR FILING DATE: 2002-10-25  
; PRIOR APPLICATION NUMBER: 60/402,342  
; PRIOR FILING DATE: 2002-08-08  
; PRIOR APPLICATION NUMBER: 10/066,390  
; PRIOR FILING DATE: 2002-02-01  
; PRIOR APPLICATION NUMBER: 60/268,785  
; PRIOR FILING DATE: 2001-02-14  
; PRIOR APPLICATION NUMBER: 60/266,386  
; PRIOR FILING DATE: 2002-02-02  
; NUMBER OF SEQ ID NOS: 35  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 32  
; LENGTH: 3637  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Encodes Cycle 3 BPP gene  
US-10-280-913A-32

Query March 72.8%; Score 2665.4; DB 17; Length 3637;  
Best Local Similarity 83.7%; Pred. No. 0;  
Matches 3065; Conservative 0; Mismatches 571; Indels 27; Gaps 3;

Qy 1 CTAATATGTAAGCCTTATATATTTGTTAAATTCGCTTAAATTTGTTAAATCAGCTC 60  
Db 3637 CTAATATGTAAGCCTTATATATTTGTTAAATTCGCTTAAATTTGTTAAATCAGCTC 3578  
Qy 61 ATTTTAAACCAATAGGCGGAAATCGGCAAAATCCCTTATATATCAAAAGATAGACCGA 120  
Db 3577 ATTTTAAACCAATAGGCGGAAATCGGCAAAATCCCTTATATCAAAAGATAGACCGA 3518  
Qy 121 GATGGGTGATGATGTTGTTCCAGTTTGGAAACAAGTCCCTATTAAGAAAGTGAAGTCTC 180  
Db 3517 GATGGGTGATGATGTTGTTCCAGTTTGGAAACAAGTCCCTATTAAGAAAGTGAAGTCTC 3458  
Qy 181 CAAGTCAAAAGGCGGAAACCGTCTATAGGGGCGATGAGCCACTAGTGAACATCAACC 240  
Db 3457 CAAGTCAAAAGGCGGAAACCGTCTATAGGGGCGATGAGCCACTAGTGAACATCAACC 3398  
Qy 241 CTAATCAAGTTTGTGGGTCGAGTGCCTTAAAGCACTAATCGAAACCTTAAAGGAG 300  
Db 3397 CTAATCAAGTTTGTGGGTCGAGTGCCTTAAAGCACTAATCGAAACCTTAAAGGAG 3338  
Qy 301 CCCCCGATTAAGCTTGAAGGGGGAAGCCGCGCAAGTGGCGGAAGAAAGGAGAA 360  
Db 3337 CCCCCGATTAAGCTTGAAGGGGGAAGCCGCGCAAGTGGCGGAAGAAAGGAGAA 3278  
Qy 361 AGCGAAAGAGCGGCGCTAGGGCGCTGGCAAGTGAAGGCTGACGCTGCGCTGAACAC 420  
Db 3277 AGCGAAAGAGCGGCGCTAGGGCGCTGGCAAGTGAAGGCTGACGCTGCGCTGAACAC 3218  
Qy 421 CACACCGGCGGCTTATAGCGCGCTACAGGGGCGCTCCATTTGCTTACAGGCTGCG 480  
Db 3217 CACACCGGCGGCTTATAGCGCGCTACAGGGGCGCTCCATTTGCTTACAGGCTGCG 3158  
Qy 481 CAACTGTGGGAAGGGGGAATCGTGCGGCGCTCTTCCGCTATTAAGCGCAAGTGGGAAAG 540  
Db 3157 CAACTGTGGGAAGGGGGAATCGTGCGGCGCTCTTCCGCTATTAAGCGCAAGTGGGAAAG 3098  
Qy 541 GGGATGTGTGCAAGGCGATTAAGTGGGTGAACGCAAGGCTTTCCAGTCAACGACGTTG 600  
Db 3097 GGGATGTGTGCAAGGCGATTAAGTGGGTGAACGCAAGGCTTTCCAGTCAACGACGTTG 3038



QY 601 TAAACGACGCGCACTGACGCGCGCTGCTTCATT CAGCTTTTGAACCGGTGAGACG 660  
Db 3037 TAAACGACGCGCACTGACGCGCGCTGCTTCATT CAGCTTTTGAACCGGTGAGACG 660  
QY 661 GGAAGCTGCGGTGCAATGTGTTTACAGCTGATGAGCAGATGAAATGCTGCACAC 720  
Db 2977 CCGCGGTGCGCGCTCTGAACTAGTGATCCCCCGGCTGCGAAATCTTATTGT 2918  
QY 721 GCTCAGAACACGAGCTAATTAACCTTAAGAAAGATTAATTAATGAGCTTA 780  
Db 2917 ATAGTATCATCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780  
QY 781 AGATATTC -ATGCTAAATTAAGCATGGAATCTGTAATACAGCTCACTAATGAGCGAA 839  
Db 2857 GGTACGCTTTTGTGTTGGGATCTTTGAAAGGCGCAATTTGTGCAAGTAAATGTGT 2798  
QY 840 TTGGGTACGCGGCGCGCGCTGCAAGGTGCAAGGTATGATTAAGCTTGAATGAAATTCG 899  
Db 2797 CTGTTAAAGACAGGCGCATGCGCAATTTGAGATTTTGTGATTAATGATGATGAT 2738  
QY 900 CAGCCCGGGAATCCATAGTCTAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCTT 959  
Db 2737 GAAAGGATCATCTTCAATGTTGTGCGAATTTTGAAGTTAGCTTGAATTCATCTTT 2678  
QY 960 TTCCCTTAAGTAAAGGTTAATTAATCCATGCTCAATTTTACGAGACTATCTTTA 1019  
Db 2677 GTTGTGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2618  
QY 1020 GGGTTAATCTAGCTGATCAGAGATCATATGCTGCGGCTTTTTCGCGCTCATGATCGC 2618  
Db 2617 GATGTTTCATCTTCTTAAATCAATACCTTTTAACTGATC-----GATTAACAG 2563  
QY 1080 CCAAGCTGCGCTATCTGCGCATCGGAGAGAAAGCGCGCTTTTCCGCGAGGT 1139  
Db 2562 GGTATCACTTCAAACTTGACTTCAAGCGCTTTTGAATTCCTGATCTTTGAAGA 2503  
QY 1140 GAAACGCGCATGAAAGATTTGCGAGATGATGATGATGATGATGATGATGATGAT 1199  
Db 2502 TATAGTGTCTCTGATCAATCTTCCGCGCATGCACTTGAAGATGATGATGATGAT 2443  
QY 1200 CGCAGCTTAACTAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1259  
Db 2442 CATATGATCGGATTAAGAGAAACATTTGAACACCATGAGAGAAATGATGATGATGAT 2383  
QY 1260 TTGCTTCAAGCGCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1319  
Db 2382 TGGCATGGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2323  
QY 1320 GTCAAGCGGAGCGCATGAGCAACCGAACAATCCGCGCAAGCGGAACTGCGGTGC 1379  
Db 2322 TGTAGCATCACTTCACTCTTCACTGACAGAAATTTGTGCCATTTAACTACATC 2263  
QY 1380 GGTGTGAGATTAATGACAGCGGTGCGCGCTGGAATTAAGTCAAGGAGAGCGGTAT 1439  
Db 2262 TAATTAACAAGAT-----TGGACAACCTCCAGTGAAGATTC 2224  
QY 1440 CTTGCGTGTGATGCGCGCAAAATGAGATGATGATGATGATGATGATGATGATGAT 1499  
Db 2223 TTCTCTTATCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2164  
QY 1500 TGGGTAAATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1559  
Db 2163 TGGGTAAATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2104  
QY 1560 ACAACATGAGCGCGGAGCATTAAGTAAAGCTGCGGAGCTTAATGATGATGATGAT 1619  
Db 2103 ACAACATGAGCGCGGAGCATTAAGTAAAGCTGCGGAGCTTAATGATGATGATGAT 2044  
QY 1620 TCAATTAATGCTTGTGCTCACTGCGCGCTTTTCACTGCGGAAACTGCTGCGCAC 1679  
Db 2043 TCAATTAATGCTTGTGCTCACTGCGCGCTTTTCACTGCGGAAACTGCTGCGCAC 1984

QY 1680 TGCATTAATGATCGGCCAAGCGCGGAGAAAGCGGTTTCCATATTTGGCGCTCTTCG 1739  
Db 1983 TGCATTAATGATCGGCCAAGCGCGGAGAAAGCGGTTTCCATATTTGGCGCTCTTCG 1924  
QY 1740 CTTCTGCTCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1799  
Db 1923 CTTCTGCTCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1864  
QY 1800 ACTCAAGCGGTAAATCGTTATCAAGATCAGGAGTAAACGAGAAAGATCATGT 1859  
Db 1863 ACTCAAGCGGTAAATCGTTATCAAGATCAGGAGTAAACGAGAAAGATCATGT 1804  
QY 1860 GAGCAAAAGCGCAAAAGCGCAAGCGGTAAAGCGCGCTGCTGCGCTTTTCC 1919  
Db 1803 GAGCAAAAGCGCAAAAGCGCAAGCGGTAAAGCGCGCTGCTGCGCTTTTCC 1744  
QY 1920 ATAGGCTCGCGCGCGCTGACAGATCAAAATTAAGCTCAAGTCAAGTGGCGAA 1979  
Db 1743 ATAGGCTCGCGCGCGCTGACAGATCAAAATTAAGCTCAAGTCAAGTGGCGAA 1684  
QY 1980 ACCGACAGGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2039  
Db 1683 ACCGACAGGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1624  
QY 2040 CTTGCTCGCGCGCGCTTACGAGATCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 2099  
Db 1623 CTTGCTCGCGCGCGCTTACGAGATCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 1564  
QY 2100 CCGTTTCTATAGCTCAAGCTGATGATGATGATGATGATGATGATGATGATGATGAT 2159  
Db 1563 CCGTTTCTATAGCTCAAGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1504  
QY 2160 TGGGTGTGAGAGAAAGCGCGCTTACGAGATCAAGTCAAGTCAAGTCAAGTCAAGT 2219  
Db 1503 TGGGTGTGAGAGAAAGCGCGCTTACGAGATCAAGTCAAGTCAAGTCAAGTCAAGT 1444  
QY 2220 GTCTTGAATCAAGCGGTAAAGCAAGCTTATGCGCATGCGACAGCACTGATTA 2279  
Db 1443 GTCTTGAATCAAGCGGTAAAGCAAGCTTATGCGCATGCGACAGCACTGATTA 1384  
QY 2280 GATTAAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2339  
Db 1383 GATTAAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2324  
QY 2340 ACGGCTCACTGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 2399  
Db 1323 ACGGCTCACTGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1264  
QY 2400 GAAAAAGGTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2459  
Db 1263 GAAAAAGGTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1204  
QY 2460 TTGTTGAAACAGAGATTAAGCGGAGAAAAAGATCTCAAGAGATCTTTGATCT 2519  
Db 1203 TTGTTGAAACAGAGATTAAGCGGAGAAAAAGATCTCAAGAGATCTTTGATCT 1144  
QY 2520 TTTCTACGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2579  
Db 1143 TTTCTACGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1084  
QY 2580 GATTAACAAAAAGATCTTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 2639  
Db 1083 GATTAACAAAAAGATCTTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 1024  
QY 2640 TCTAAAGTATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2699  
Db 1023 TCTAAAGTATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 964  
QY 2700 CTAATCAAGGATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2759  
Db 963 CTAATCAAGGATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 904  
QY 2760 TAACTAAGATGAGAGGCTTACATCTGCGCCCAAGTGTGATGATGATGATGATGAT 2819



Db 903 TAACTACATACGAGGAGGCTTACATCTGGCCCAAGTCTGCAATGATACCGCAAGCC 844  
Oy 2820 CACGCTCACGCGCTCCAGATTATATACAGAAATAAACAGCAGCCGGAAGGCGGAGCCA 2879  
Db 843 CACGCTCACGCGCTCCAGATTATATACAGAAATAAACAGCAGCCGGAAGGCGGAGCCA 784  
Oy 2880 GAAAGGTCTCTGCACTTTATCCGCTTCATCTGATTTATTTATTTGTCGGGGAAGCTA 2939  
Db 783 GAAAGGTCTCTGCACTTTATCCGCTTCATCTGATTTATTTATTTGTCGGGGAAGCTA 724  
Oy 2940 GAGTAAGTAGTTCGCGCAGTTAATAGTTGGCGAAGTGGTGGCCATGCTACAGGATCG 2999  
Db 723 GAGTAAGTAGTTCGCGCAGTTAATAGTTGGCGAAGTGGTGGCCATGCTACAGGATCG 664  
Oy 3000 TGGTGTCAAGCTCGTCTGTTGGTATAGCTTCAATTCAGCTCCGGTCCCAACAGATCAAGC 3059  
Db 663 TGGTGTCAAGCTCGTCTGTTGGTATAGCTTCAATTCAGCTCCGGTCCCAACAGATCAAGC 604  
Oy 3060 GAGTTACATGATCCCGCATGTTGTCGCAAAAAAGCGGTTAGCTCCTTCGGTCTCCGATCG 3119  
Db 603 GAGTTACATGATCCCGCATGTTGTCGCAAAAAAGCGGTTAGCTCCTTCGGTCTCCGATCG 544  
Oy 3120 TTGTCAGAAATAGTTGGCCGAGGTTATCATCATGTTATAGCAGACATCATATTT 3179  
Db 543 TTGTCAGAAATAGTTGGCCGAGGTTATCATCATGTTATAGCAGACATCATATTT 484  
Oy 3180 CTCTTACTGTATGTCATCCGATCCGATAGATCTTTCTGTGATGAGTACTCAACCAAGT 3239  
Db 483 CTCTTACTGTATGTCATCCGATCCGATAGATCTTTCTGTGATGAGTACTCAACCAAGT 424  
Oy 3240 CATCTGGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCTCAATACGGATTA 3299  
Db 423 CATCTGGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCTCAATACGGATTA 364  
Oy 3300 ATACCGGCGCATATGAGAACTTTAAAGTCTCATATGAGAAAGGTTCTTCGGGAGC 3359  
Db 363 ATACCGGCGCATATGAGAACTTTAAAGTCTCATATGAGAAAGGTTCTTCGGGAGC 304  
Oy 3360 GAAAACTCTCAAGATCTTACCGCTGTGAGATCCAGTTCGATGTAACCACTCGTGAC 3419  
Db 303 GAAAACTCTCAAGATCTTACCGCTGTGAGATCCAGTTCGATGTAACCACTCGTGAC 244  
Oy 3420 CCAACTGATCTTCAGCATCTTTTACTTTCACCAAGGCTTCTGGGTGAGCAAAACAGAA 3479  
Db 243 CCAACTGATCTTCAGCATCTTTTACTTTCACCAAGGCTTCTGGGTGAGCAAAACAGAA 184  
Oy 3480 GGGAAAAATGCGGCAAAAAGGAAATAGGGGCGACACGGAATGTTGAATCTCATCTCT 3539  
Db 183 GGGAAAAATGCGGCAAAAAGGAAATAGGGGCGACACGGAATGTTGAATCTCATCTCT 124  
Oy 3540 TCCCTTTTCAATATATATTAAGACATTTATCAAGGTTATTTGTTCTCATGAGCGGATCATAT 3599  
Db 123 TCCCTTTTCAATATATATTAAGACATTTATCAAGGTTATTTGTTCTCATGAGCGGATCATAT 64  
Oy 3600 TTGAATGATTTAGAAAAATTAACCAATAGGGGTTCCGCGCACATTTCCCGAAAAATGC 3659  
Db 63 TTGAATGATTTAGAAAAATTAACCAATAGGGGTTCCGCGCACATTTCCCGAAAAATGC 4  
Oy 3660 CAC 3662  
Db 3 CAC 1

;; CURRENT FILING DATE: 2003-10-10  
;; PRIOR APPLICATION NUMBER: 60/402,342  
;; PRIOR FILING DATE: 2002-08-08  
;; PRIOR APPLICATION NUMBER: 10/066,390  
;; PRIOR FILING DATE: 2002-02-01  
;; PRIOR APPLICATION NUMBER: 60/266,785  
;; PRIOR FILING DATE: 2001-02-14  
;; PRIOR APPLICATION NUMBER: 60/266,386  
;; PRIOR FILING DATE: 2002-02-02  
;; NUMBER OF SEQ ID NOS: 35  
;; SOFTWARE: PatentIn version 3.1  
;; SEQ ID NO 32  
;; LENGTH: 3637  
;; TYPE: DNA  
;; ORGANISM: Artificial Sequence  
;; FEATURE:  
;; OTHER INFORMATION: Encodes Cycle 3 BFP gene  
US-10-684-134-32  
  
Query Match 72.8%; Score 2665.4; DB 17; Length 3637;  
Best Local Similarity 83.7%; Pred. No. 0;  
Matches 3065; Conservative 0; Mismatches 571; Indels 27; Gaps 3;  
  
Oy 1 CTAATGTAGAGCGTTAATATTTGTTAAATTCGCGTTAAATTTGTTAAATCAGCTC 60  
Db 3637 CTAATGTAGAGCGTTAATATTTGTTAAATTCGCGTTAAATTTGTTAAATCAGCTC 3578  
Oy 61 ATTTTAAACCAATAGCCGCAATTCGCAAAATCCCTTAATCAAAAGATAGACCGA 120  
Db 3577 ATTTTAAACCAATAGCCGCAATTCGCAAAATCCCTTAATCAAAAGATAGACCGA 3518  
Oy 121 GATAGGGTGAAGTGTGTTCCAGTTTGGAAACAAGTTCACATTTAAAGAACTGAGCTC 180  
Db 3517 GATAGGGTGAAGTGTGTTCCAGTTTGGAAACAAGTTCACATTTAAAGAACTGAGCTC 3458  
Oy 181 CAAGTCAAGAGGGGGAAGAAACCGTCTATCAAGGGGAGTGGCCCATAGTGAACCATCAC 240  
Db 3457 CAAGTCAAGAGGGGGAAGAAACCGTCTATCAAGGGGAGTGGCCCATAGTGAACCATCAC 3398  
Oy 241 CTAATCAAGTTTGTGGGCTGAGAGTCCGTTAAAGCACTAAATCGGAACCTTAAGGAG 300  
Db 3397 CTAATCAAGTTTGTGGGCTGAGAGTCCGTTAAAGCACTAAATCGGAACCTTAAGGAG 3338  
Oy 301 CCCCCGATTTAGACTTGAACGGGGAAGCCGGCGAAGCTGCGGAAGAAAGGAAAGAA 360  
Db 3337 CCCCCGATTTAGACTTGAACGGGGAAGCCGGCGAAGCTGCGGAAGAAAGGAAAGAA 3278  
Oy 361 AGCGAAGAGAGCGGCGCTAGGGGCTGGGCAAGTGTAGCGGTCAAGCTGCGGCTAACAC 420  
Db 3277 AGCGAAGAGAGCGGCGCTAGGGGCTGGGCAAGTGTAGCGGTCAAGCTGCGGCTAACAC 3218  
Oy 421 CACACCGCGCGCTTAATGCGCGCTACAGAGGCGCTCCATTCGCGCATTCAGGCTCGC 480  
Db 3217 CACACCGCGCGCTTAATGCGCGCTACAGAGGCGCTCCATTCGCGCATTCAGGCTCGC 3158  
Oy 481 CAACGTGTGGAGAGGCGATGAGTGGCGGCTCTTTCGCTATTACGCCAGCTGGGAAAG 540  
Db 3157 CAACGTGTGGAGAGGCGATGAGTGGCGGCTCTTTCGCTATTACGCCAGCTGGGAAAG 3098  
Oy 541 GGGATGTGTCGCAAGCGAATTAAGTTGGGTAAAGCGCAGGCTTTCCAGTCAAGAGTGG 600  
Db 3097 GGGATGTGTCGCAAGCGAATTAAGTTGGGTAAAGCGCAGGCTTTCCAGTCAAGAGTGG 3038  
Oy 601 TAAACGACGCGCAGTAGAGCGCGCTGTTCAATCAAGTTTGAACCCGTTGAGAGCG 660  
Db 3037 TAAACGACGCGCAGTAGAGCGCGCTGTTCAATCAAGTTTGAACCCGTTGAGAGCTCA 2978  
Oy 661 GCAGACTGCGGTCGCAATGTGTTTAAACGCTGATGAGAGCAATGATGCTCGACAC 720  
Db 2977 CCGGCGGTGCGCGCTCTAGAACTAGATGATCCCGGCGCTGACGAAATCTTAATTTGT 2918  
Oy 721 GCTCAGAACGAGCTAGATTAAACCTTGAAGAAATTAATCAATTTGTGACGTAGTTAA 780

RESULT 5  
US-10-684-134-32/c  
; Sequence 32, Application US/10684134  
; Publication No. US20040142433A1  
; GENERAL INFORMATION:  
; APPLICANT: Large Scale Biology Corporation  
; TITLE OF INVENTION: POPULATION OF POLYNUCLEOTIDE SEQUENCE VARIANTS  
; FILE REFERENCE: LSBC-PLG5192-CIP  
; CURRENT APPLICATION NUMBER: US/10/684,134

Db 2917 ATAGTCAATCATCCATGCTATGTAATCCAGCAGAGTTACAAACTCAAGAAAGAACATGT 2858  
 QY 781 AGATAATC-ATGGGTAAATTTGACGCATGGGATCTGTATATACACTCATATAGGCGAA 839  
 Db 2857 GGTACGCTTTTGTGGGATCTTTGAAAGGCGAGATTTGTGTGACAGTTAATGGTTGT 2798  
 QY 840 TTGGGTACCGGGCCCCCTCGAGGTCCAGGTATGATTAAGCTTATATGAATTCCTG 899  
 Db 2797 CTGGTAAAGGACAGGCGCATTCGCCAATTTGAGTATTTTGTGATTAATGTCTGTAGT 2738  
 QY 900 CAGCCCGGGGATCCACTAGTTCATAGACCGCGCCACCGCGGTGAGCTCCAGCTTTG 959  
 Db 2737 GAAAGGATCCATCTTCAATGTGTGGCAATTTTGAAGTTAGCTTTGATTCATCTTTT 2678  
 QY 960 TTCCCTTATAGTAGGGTTAATTAATGCCATGCGCTCAATTTTACGACATCTTTCTA 1019  
 Db 2677 GTTGTCTCGCGTATGTAATGTAATTTGTGTGATTAAGTTGATCTCAAGTTGTGTCGA 2618  
 QY 1020 GGGTAAATCTAGCTGCATCAGGATCATATCGTGGGTCTTTTTCGGGCTCACTATCCG 1079  
 Db 2617 GAATGTTTCATCTTCTTAAATCAATACCTTTTATCTGATAC-----GATTAACAG 2563  
 QY 1080 CCAAGCTGGCGCTTATCGGGCATGGGGAGAAAGAACCCGTCTTTTCCGCGAGATT 1139  
 Db 2562 GGTATCACTTCAAACTTGACTTACGACGCGTCTTGTAGTTCCGTCATCTTTGAAAG 2503  
 QY 1140 GAACGGCATGAAAGAGTTTCCGAGATGATCTGCTGCTGATTAAGCTTGAAGGAAA 1199  
 Db 2502 TATAGTCTGCTCTGTATCATATACCTTCGGGCAATGGCACTTGAAGAAAGTCATGCGTTT 2443  
 QY 1200 CGCAGCTTTACATGATGATTGCGGAAGGTGTGCGCATGACAGCTTTTAAACGTTAACTG 1259  
 Db 2442 CATATGATCGGATTAACGAGAAAAGCATGAAACCATGAGAGAAAGTAGTACAAAGT 2383  
 QY 1260 TTGCTTACGGCCACTGGGATACCAAGTGTGCGGCTTTTCCGACACATTCGAGT 1319  
 Db 2382 TGCCCATGAGAAAGTAGTTTTCAGTAGTGCATAATTTAATGTAACTTTCCGTA 2323  
 QY 1320 GTACGCCGGAAGGCGATCAGCAACCCGAACTACCGCGACAGCGGAACTGCGCTGCC 1379  
 Db 2322 TGTAGATCACTTCACTTCCCTTCCACTGACAGAAAATTTGTGCCATTTAATCATCAGT 2263  
 QY 1380 GGTGTCAATTAATGACAGCGGTGCGGCTGGGATTAATGATGACAGAGAGGTAT 1439  
 Db 2262 TAATTCACAGAAAT-----TGGACAACTCAGTGAAGAAAGTTC 2224  
 QY 1440 CCTGCTGATGTCGCGAGAAATGACATGATACCCCGTGAATTAACCGCGGCGCGCT 1499  
 Db 2223 TTCTCTTACTCATCGTACCAGCTTTGTCTCTTATGAGGGTTAATTTGCCGCT 2164  
 QY 1500 TGGCGTATCATGCTCATAGCTGTTTCTGTGTAATTTGTTATCCGCTCACAATTCAC 1559  
 Db 2163 TGGGTATCATGCTCATAGCTGTTTCTGTGTAATTTGTTATCCGCTCACAATTCAC 2104  
 QY 1560 ACAATATCAGAGCGGAGCAATAAGCTTAAGCTTGGGTGCTAATGATGAGCTAAC 1619  
 Db 2103 ACAATATCAGAGCGGAGCAATAAGCTTGGGTGCTAATGATGAGCTAAC 2044  
 QY 1620 TCAATTAATTTGGCTTGGCTCACTGCGCTTTTCCAGTGGGAAACCTGTGCTGAC 1679  
 Db 2043 TCAATTAATTTGGCTTGGCTCACTGCGCTTTTCCAGTGGGAAACCTGTGCTGAC 1984  
 QY 1680 TGCATTAATGATTCGGCAACGGCGGGAGAGCGGTTTGGCTAATTTGGGCGCTTTCCG 1739  
 Db 1983 TGCATTAATGATTCGGCAACGGCGGGAGAGCGGTTTGGCTAATTTGGGCGCTTTCCG 1924  
 QY 1740 CTTCCTGCTCACTGATCTGCTGCGCTCGCTGCTTGGCTGCGGAGACCGTATCAGCTC 1799  
 Db 1923 CTTCCTGCTCACTGATCTGCTGCGCTCGCTGCTTGGCTGCGGAGACCGTATCAGCTC 1864  
 QY 1800 ACTCAAGGCGGTAATACGTTATATCAAGATCAAGGAGATTAAGCAAGAAAGCAATG 1859  
 Db 1863 ACTCAAGGCGGTAATACGTTATATCAAGATCAAGGAGATTAAGCAAGAAAGCAATG 1804

QY 1860 GAGCAAAAGCCAGCAAAAGCCAGCAACCGTAATAAGCCGCGTGTGCTGCGCTTTTCC 1919  
 Db 1803 GAGCAAAAGCCAGCAAAAGCCAGCAACCGTAATAAGCCGCGTGTGCTGCGCTTTTCC 1744  
 QY 1920 ATAGGCTCGCCCCCTGACAGAGCATACAAAATATGAGCTCAATCAGAGTGGCGAA 1979  
 Db 1743 ATAGGCTCGCCCCCTGACAGAGCATACAAAATATGAGCTCAATCAGAGTGGCGAA 1684  
 QY 1980 ACCGACAGAGCTAATAAGTATCCAGGCGTTTCCCCCTGAAAGCTCCCTGTGCTCTC 2039  
 Db 1683 ACCGACAGAGCTAATAAGTATCCAGGCGTTTCCCCCTGAAAGCTCCCTGTGCTCTC 1624  
 QY 2040 CTGTTCCGACCTTCCGCTTACCGGATACCTGTGCGCTTCTCCCTTCCGAAAGCTGG 2099  
 Db 1623 CTGTTCCGACCTTCCGCTTACCGGATACCTGTGCGCTTCTCCCTTCCGAAAGCTGG 1564  
 QY 2100 CGCTTCTCATACCTGATGAGTATCTCAATTCGAGTGTGAGTGTGCTGCTCAAGC 2159  
 Db 1563 CGCTTCTCATACCTGATGAGTATCTCAATTCGAGTGTGAGTGTGCTGCTCAAGC 1504  
 QY 2160 TGGCTGTGTGACGAAACCCCGTTCAGCCCGACCGCTGGGCTTATCCGTTAATATC 2219  
 Db 1503 TGGCTGTGTGACGAAACCCCGTTCAGCCCGACCGCTGGGCTTATCCGTTAATATC 1444  
 QY 2220 GTCTGAGTCCAAACCGGTAAGACAGACTTATGCGCATGCGCAGACCACTGTATAC 2279  
 Db 1443 GTCTGAGTCCAAACCGGTAAGACAGACTTATGCGCATGCGCAGACCACTGTATAC 1384  
 QY 2280 GGAATTAAGCAGAGGATGATGAGCGGTGTCTAAGAGTTCTTGAATGAGTGGCTTAACT 2339  
 Db 1383 GGAATTAAGCAGAGGATGATGAGCGGTGTCTAAGAGTTCTTGAATGAGTGGCTTAACT 1324  
 QY 2340 ACGGCTACACTAAGAGGACATATTTGTATCTGCGCTGTCTGTAAGCAAGTTACTTGC 2399  
 Db 1323 ACGGCTACACTAAGAGGACATATTTGTATCTGCGCTGTCTGTAAGCAAGTTACTTGC 1264  
 QY 2400 GAAAAAGTTGTAGTCTTGTATCCGGCAAAACAAACCGGTGTAGCGGTGTTTTT 2459  
 Db 1263 GAAAAAGTTGTAGTCTTGTATCCGGCAAAACAAACCGGTGTAGCGGTGTTTTT 1204  
 QY 2460 TTGTTTCAAGCAGCAGATTAACCGCGAATAAAGATCTCAAGAGATCTTGTATCT 2519  
 Db 1203 TTGTTTCAAGCAGCAGATTAACCGCGAATAAAGATCTCAAGAGATCTTGTATCT 1144  
 QY 2520 TTCTACGGGGTGTGACGCTCACTGGAAGAAACCTCAAGTTAAGGATTTTGTATGA 2579  
 Db 1143 TTCTACGGGGTGTGACGCTCACTGGAAGAAACCTCAAGTTAAGGATTTTGTATGA 1084  
 QY 2580 GATTATCAAAAAGATCTTCACTGATATCTTTTAAATTAATAAGATTTTAAATCA 2639  
 Db 1083 GATTATCAAAAAGATCTTCACTGATATCTTTTAAATTAATAAGATTTTAAATCA 1024  
 QY 2640 TCTTAAGTATATAGATACTTAAGTGTGACAGTTTACCAATGCTTAAATCAGTAGGAC 2699  
 Db 1023 TCTTAAGTATATAGATACTTAACTTGTGACAGTTTACCAATGCTTAAATCAGTAGGAC 964  
 QY 2700 CTATCTACAGCATCTGTCTAATTTGTTGATCATATGTTGCTGATCTCCCGTGTGAG 2759  
 Db 963 CTATCTACAGCATCTGTCTAATTTGTTGATCATATGTTGCTGATCTCCCGTGTGAG 904  
 QY 2760 TAACTACGATACGAGAGGCTTACCATCTGCGCCCAATGCTGCAATGATACCGAGACC 2819  
 Db 903 TAACTACGATACGAGAGGCTTACCATCTGCGCCCAATGCTGCAATGATACCGAGACC 844  
 QY 2820 CACGCTCACCGGCTTCAGATTTATACCAATAAACAAGCAGCGGAGAGGCGGAGCGCA 2879  
 Db 843 CACGCTCACCGGCTTCAGATTTATACCAATAAACAAGCAGCGGAGAGGCGGAGCGCA 784  
 QY 2880 GAAAGTGTCTCGCAACTTTTATCCGCTCCATCAAGTCTATTAATTTTGGGAAAGCTTA 2939  
 Db 783 GAAAGTGTCTCGCAACTTTTATCCGCTCCATCAAGTCTATTAATTTTGGGAAAGCTTA 724

Qy	2940	NAGTAAAGAGTATCGCCACATTAATATGTTGGCCAGACGTTGGTCCATTGCTACAGGCATCG	2939
Db	723	GAGTAAGAGTATCGCAGATTAATATGTTGGCCAGACGTTGGTCCATTGCTACAGGCATCG	664
Qy	3000	TGATGTACGCGCTCGCTTGTGGTATAGGCTTCATTCAGGCTCCGGTATCCCAAGATCAAGGC	3059
Db	663	TGGTGTACGCTCGCTCGCTTGTGGTATAGGCTTCATTCAGGCTCCGGTATCCCAAGATCAAGGC	604
Qy	3060	GAGTTACATGATATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCTTCGGTCCCTCGATCG	3119
Db	603	GAGTTACATGATATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCTTCGGTCCCTCGATCG	544
Qy	3120	TTGTTCAGAAAGTAAGTTGGCCGCGCAGTGTATCATCTATAGGTTATGGCAGCATCTGATAATT	3179
Db	543	TTGTTCAGAAAGTAAGTTGGCCGCGCAGTGTATCATCTATAGGTTATGGCAGCATCTGATAATT	484
Qy	3180	CTCTTACCTGATGTCATCGGTAAGATGCTTTTCGTGATCAGTGATGATCAACCAAGT	3239
Db	483	CTCTTACCTGATGTCATCGGTAAGATGCTTTTCGTGATCAGTGATGATCAACCAAGT	424
Qy	3240	CATTCTGAAATAGTGTATGTCGGCGCACCGAGTTGCTCTTGGCCGCGCTCAATACGGGATA	3299
Db	423	CATTCTGAAATAGTGTATGTCGGCGCACCGAGTTGCTCTTGGCCGCGCTCAATACGGGATA	364
Qy	3300	ATACCGCGCCACATACGCAAACTTTAAAGTGTCTATCATTTGAAAAAGCTTCTTGCGGGC	3359
Db	363	ATACCGCGCCACATACGCAAACTTTAAAGTGTCTATCATTTGAAAAAGCTTCTTGCGGGC	304
Qy	3360	GAAAACTCTCAAGGATCTTACCGCTGTGATCCAGTTGCATATGATCAATCCCATCTGTGCAC	3419
Db	303	GAAAACTCTCAAGGATCTTACCGCTGTGATCCAGTTGCATATGATCAATCCCATCTGTGCAC	244
Qy	3420	CCAACGTATCTTCAGCATCTTTTACTTTCACACAGCGTTCTGGGTGACCAAAAAACAGGA	3479
Db	243	CCAACGTATCTTCAGCATCTTTTACTTTCACACAGCGTTCTGGGTGACCAAAAAACAGGA	184
Qy	3480	GGCAAAATGCGCGCAAAAAAGGAAATAGAGGGCGACACGGAAATTTGAATACTACTCT	3539
Db	183	GGCAAAATGCGCGCAAAAAAGGAAATAGAGGGCGACACGGAAATTTGAATACTACTCT	124
Qy	3540	TCCTTTTTCATATTTATGGAAGCATTTATCAGGGTATTTGTCTCATAGCGCGATACATAT	3599
Db	123	TCCTTTTTCATATTTATGGAAGCATTTATCAGGGTATTTGTCTCATAGCGCGATACATAT	64
Qy	3600	TTGAATGATTTAGAAAAATAAACAAATAGGGGTTCCGCGACATTTCCCGAAAAAGTGC	3659
Db	63	TTGAATGATTTAGAAAAATAAACAAATAGGGGTTCCGCGACATTTCCCGAAAAAGTGC	4
Qy	3660	CAC 3662	
Db	3	CAC 1	
RESULT 6			
US-10-637-758-32/c			
Sequence 32, Application US/10637758			
Publication No. US20040180352A1			
GENERAL INFORMATION:			
APPLICANT: Large Scale Biology Corporation			
TITLE OF INVENTION: POPULATION OF POLYNUCLEOTIDE SEQUENCE VARIANTS			
FILE REFERENCE: LSBC-PI65392-CIP			
CURRENT APPLICATION NUMBER: US/10/637,758			
CURRENT FILING DATE: 2003-08-08			
PRIORITY APPLICATION NUMBER: US/10/280,913			
PRIORITY FILING DATE: 2002-10-25			
PRIORITY APPLICATION NUMBER: 60/402,342			
PRIORITY FILING DATE: 2002-08-08			
PRIORITY APPLICATION NUMBER: 10/066,390			
PRIORITY FILING DATE: 2002-02-01			
PRIORITY APPLICATION NUMBER: 60/268,785			
PRIORITY FILING DATE: 2001-02-14			
PRIORITY APPLICATION NUMBER: 60/266,386			
PRIORITY FILING DATE: 2002-02-02			

	Query Match	Beef Local Similarity	Score	DB	Length
	72.8%;	83.7%;	2665.4;	17;	3637;
	Matches 3065;	Conservative 0;	Mismatches 571;	Indels 27;	Gaps 3;
	US-10-637-758-32				
	OTHER INFORMATION: Encodes Cycle 3 BFP gene				
	PEATURE:				
	ORGANISM: Artificial Sequence				
	TYPE: DNA				
	LENGTH: 3637				
	SEQ ID NO 32				
	SOFTWARE: PatentIn version 3.1				
	NUMBER OF SEQ ID NOS: 35				
QY	1	CTAAATTGAAGCGTTAATATTTTGT	AAAATTCGCGTTAAATTTTGT	AAATCAGCTC	60
DB	3637	CTAAATTGAAGCGTTAATATTTTGT	AAAATTCGCGTTAAATTTTGT	AAATCAGCTC	3578
QY	61	ATTTTAAACCAATAGCCGGAATCG	CAAAATCCCTTAAATCAAAAGAA	TAGACCGA	120
DB	3577	ATTTTAAACCAATAGCCGGAATCG	CAAAATCCCTTAAATCAAAAGAA	TAGACCGA	3518
QY	121	GATGGGGTGAAGTGTGTTCCAGT	TGGAAACAAGTCCATATTAAGA	ACGTGACTC	180
DB	3517	GATGGGGTGAAGTGTGTTCCAGT	TGGAAACAAGTCCATATTAAGA	ACGTGACTC	3458
QY	181	CAACGTCAAAGGGCGAATAACG	TCATCAGGGGATGACCCTAC	TGTAACATCAC	240
DB	3457	CAACGTCAAAGGGCGAATAACG	TCATCAGGGGATGACCCTAC	TGTAACATCAC	3398
QY	241	CTAATCAAGTTTTTGGGGTGA	AGTGCCTTAAGCACTAAATCG	GAACCTTAAAGGAG	300
DB	3397	CTAATCAAGTTTTTGGGGTGA	AGTGCCTTAAGCACTAAATCG	GAACCTTAAAGGAG	3338
QY	301	CCCCCGATTGAGCTTGA	CGGGGAAAGCCGGCGAA	CGTGGCGA	360
DB	3337	CCCCCGATTGAGCTTGA	CGGGGAAAGCCGGCGAA	CGTGGCGA	3278
QY	361	AGCGAAAGGACCGGGCGCTAG	GGGCGTGAAGGTAGCGGCT	CAACGCTGAC	420
DB	3277	AGCGAAAGGACCGGGCGCTAG	GGGCGTGAAGGTAGCGGCT	CAACGCTGAC	3218
QY	421	CACACCCGCGCGCTTA	TGCGCCGCTACAGGGCG	CGTCCATTCA	480
DB	3217	CACACCCGCGCGCTTA	TGCGCCGCTACAGGGCG	CGTCCATTCA	3158
QY	481	CAACTGTTGGGAAGGGCG	ATGCGTGGCGGCTCTT	CGCTATTA	540
DB	3157	CAACTGTTGGGAAGGGCG	ATGCGTGGCGGCTCTT	CGCTATTA	3098
QY	541	GGGATGTCGTCAGAGCG	ATTAAGTTGGGTAAAGCC	AGGGTTTCC	600
DB	3097	GGGATGTCGTCAGAGCG	ATTAAGTTGGGTAAAGCC	AGGGTTTCC	3038
QY	601	TAAACGACGCGCAGTGA	CGCGCGCTGCTTCA	TTCAGATTA	660
DB	3037	TAAACGACGCGCAGTGA	CGCGCGCTGCTTCA	TTCAGATTA	2978
QY	661	GCAGACTCGCGTC	CAATGTGTTTCAAGCG	TAGTGAAGATG	720
DB	2977	GCAGACTCGCGTC	CAATGTGTTTCAAGCG	TAGTGAAGATG	2918
QY	721	GCTGCAGAACGACG	AGCTAATTAACCTT	AGAAAGATAT	780
DB	2917	GCTGCAGAACGACG	AGCTAATTAACCTT	AGAAAGATAT	2858
QY	781	AGATATATC-ATG	CGTAAATTAAGCA	TGCGATCTGTA	839
DB	2857	AGATATATC-ATG	CGTAAATTAAGCA	TGCGATCTGTA	2798
QY	840	TTGGGTACCGGG	CCCCCTTCGAG	STGACGATTA	899
DB	2797	TTGGGTACCGGG	CCCCCTTCGAG	STGACGATTA	2718

QY 900 CAGCCGGGGGATCCACTGATTCTTGAAGCGCGCCGACCGGGGTGAGGCTCCAGCTTTTG 959  
DB 2737 GAGCGATCCATCTTCAATGTTGTGGCAATTTTGAAGTTAGTTGATTCATCTTTT 959  
QY 960 TTCCCTTAAATGAGGGTTAATTAATGATCCCATGGGTCAATTTTAAACGACATCTTTCTA 2678  
DB 2677 GTTGTCTGCGCGTATGATATACATGTGTGAAGTTAAAGTTGATCTGAGTTTGTGTCCA 2618  
QY 1020 GGGTTAATCTAGCTGATCAGATCAGATATCGTGGGTCTTTTTCGGCTGAGCATCCG 2618  
DB 2617 GTTGTCTGCGCGTATGATATACATGTGTGAAGTTAAAGTTGATCTGAGTTTGTGTCCA 2618  
QY 2617 GATGTTTCCATCTCTTTAAATCAATACCTTTTAAGTCATAC-----GATTAACAAG 2563  
DB 1080 CCAAGCTGGCGCTATCTGGGCAATCGGGAGGAAAGCCCGTCTTTTCCCGGAGGTT 1139  
QY 2562 GGTATCACCTTCAAACTTGAATTCAGCAAGCTCTTGAATTCCTGATCTTTGAAAGA 2503  
DB 1140 GAAGCGGAGGAGGAAAGTTTCCGAGATGATCTGCTGCTGATTTGAAGTTGAGGAGAA 1199  
QY 2502 TATAGTCTTCTCTGATCAATACCTTCGGGATGGAATCTTGAAGGATCAATGCGGTTT 2443  
DB 1200 CGCATGTTTACATGATGATTCGGGAAGTGTGGCATGCAATGAAAGTCAATGCGGTTT 2443  
QY 2442 CATATGATTCGGATTAACGAGAAAGCATTTGAACACATGAGAAAGTATGACAAAGTG 2383  
DB 1260 TTCTGTTAGGCACTGGGATACAGATTCGTGGGCTTTTCCGACACAGTTCCGAGATG 1319  
QY 2382 TGGCATGGAACAGATGATTTTCCAGATGTCAAATTAATTAAGTGAAGCTTTCCGTA 2323  
DB 1320 GTGAGCCGGAAGCCGATCAGCAACCGCAATACCGGAGACAGCGGAACTGCGGTCC 1379  
QY 2332 TGTGACATCACTTCACTCTCTCAGTACAGAAATTTGTGCCATTAACATCAATC 2263  
DB 1380 GGTGTGAGATTAATGACAGCGGTGGCGGCTGGGATTAATGATGAGGAGAGAGGAT 1439  
QY 2262 TAATTCACAGAAAT-----TGGACAACTCCAGTGAAGTTC 2224  
DB 1440 CTGCGCTGATGTCGAGAAATGACATGATACCCCGTGAATACCGGCGGCGGCT 1499  
QY 2223 TTCTCTTTAACTCATGATGATACCGATTTTGTCTTAACTGAGGTTAATGCGGCT 2164  
DB 1500 TGGCGTAAATCATGATGATGCTTTTCTGTGTGAATTTGTTATCCGCTCAATTCAC 1559  
QY 2163 TGGCGTAAATCATGATGATGCTTTTCTGTGTGAATTTGTTATCCGCTCAATTCAC 2104  
DB 1560 ACAACATTAAGAGCCGGAAGCATTAAGTTAAAGCTGGGCTTAATGATGAGCTTAAC 1619  
QY 2103 ACAACATTAAGAGCCGGAAGCATTAAGTTAAAGCTGGGCTTAATGATGAGCTTAAC 2044  
DB 1620 TCACATTAATGCGTTGCTGCTCACTGCCCTTTCCAGTGGGAAACCTGTGTGCCAGC 1679  
QY 2043 TCACATTAATGCGTTGCTGCTCACTGCCCTTTCCAGTGGGAAACCTGTGTGCCAGC 1984  
DB 1680 TGCATTAATGATGCGCAACGCGCGGGAGAGCGGTTTGGGATTTGGGCGCTTTCCG 1739  
QY 1983 TGCATTAATGATGCGCAACGCGCGGGAGAGCGGTTTGGGATTTGGGCGCTTTCCG 1924  
DB 1740 CTTCCTGCTCACTGATCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1799  
QY 1923 CTTCCTGCTCACTGATCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1864  
DB 1800 ACTCAAGGCGGTTAATCGGTTATCCAGAAATCAAGGGGATTAACGAGAAAGACATGT 1859  
QY 1863 ACTCAAGGCGGTTAATCGGTTATCCAGAAATCAAGGGGATTAACGAGAAAGACATGT 1804  
DB 1860 GAGCAAAAGGCGAGAAAGGCGAGAACCGTTAAAGGCGCGGTGTGCTGGCTTTTCC 1919  
QY 1803 GAGCAAAAGGCGAGAAAGGCGAGAACCGTTAAAGGCGCGGTGTGCTGGCTTTTCC 1744  
DB 1920 ATAGGCTCGCCCTCTGACGATCAAAATTCAGCTGAAAGTCAAGGTGCGAA 1979  
QY 1743 ATAGGCTCGCCCTCTGACGATCAAAATTCAGCTGAAAGTCAAGGTGCGAA 1684

---

QY 1980 ACCCGACAGACTTAATTAAGATACAGGCGTTTCCCTGAGAGTCCCTGTGCTTC 2039  
DB 1683 ACCCGACAGACTTAATTAAGATACAGGCGTTTCCCTGAGAGTCCCTGTGCTTC 1624  
QY 2040 CTGTTCGACCTCGCGCTTACCGGATACCTGTCCGCTTTTCCCTTGGGAGCGTGG 2099  
DB 1623 CTGTTCGACCTCGCGCTTACCGGATACCTGTCCGCTTTTCCCTTGGGAGCGTGG 1564  
QY 2100 CGCTTTCATATGCTCAAGCTGATGATCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCT 2159  
DB 1563 CGCTTTCATATGCTCAAGCTGATGATCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCT 1504  
QY 2160 TGGGCTGTGTGCAAGAACCCCGGTCAAGCCGACCCGCTGCGCTTATCCGTAATC 2219  
DB 1503 TGGGCTGTGTGCAAGAACCCCGGTCAAGCCGACCCGCTGCGCTTATCCGTAATC 1444  
QY 2220 GTCTTGAATCCACCCGTTAAGACAGACTTATCCGACTGAGAGCCGACCTGTAACA 2279  
DB 1443 GTCTTGAATCCACCCGTTAAGACAGACTTATCCGACTGAGAGCCGACCTGTAACA 1384  
QY 2280 GATTAAGAGAGGATATGATGCGGTGCTGACAGATTTTGAAGTGTGCTTAAT 2339  
DB 1383 GATTAAGAGAGGATATGATGCGGTGCTGACAGATTTTGAAGTGTGCTTAAT 1324  
QY 2340 ACGGCTACACTGAGAGGACAGATTTTGTATCTGCGCTGCTGAGAGCCGTTAATC 2399  
DB 1323 ACGGCTACACTGAGAGGACAGATTTTGTATCTGCGCTGCTGAGAGCCGTTAATC 1264  
QY 2400 GAAAAAGAGTTGATGCTTGTATCCGGAACAAACCCCTGTGAGCGGTGTTT 2459  
DB 1263 GAAAAAGAGTTGATGCTTGTATCCGGAACAAACCCCTGTGAGCGGTGTTT 1204  
QY 2460 TTGTTGCAAGCAGACATTAACCGGCAAGAAAAAGATCTCAAGAGATCTTTGATCT 2519  
DB 1203 TTGTTGCAAGCAGACATTAACCGGCAAGAAAAAGATCTCAAGAGATCTTTGATCT 1144  
QY 2520 TTGTTGCAAGCAGACATTAACCGGCAAGAAAAAGATCTCAAGAGATCTTTGATCT 2579  
DB 1143 TTGTTGCAAGCAGACATTAACCGGCAAGAAAAAGATCTCAAGAGATCTTTGATCT 1084  
QY 2580 GATTAAGAGAGTCTTCACTGATCTTTTAAATTAATTAATTAATTAATTAATTAAT 2639  
DB 1083 GATTAAGAGAGTCTTCACTGATCTTTTAAATTAATTAATTAATTAATTAATTAAT 1024  
QY 2640 TCTAAAGTATTAATGATTAATCTGATGACAGTTTAACTTAATCAATGAGCAC 2699  
DB 1023 TCTAAAGTATTAATGATTAATCTGATGACAGTTTAACTTAATCAATGAGCAC 964  
QY 2700 CTATCTCAGGATCTGATCTTATTTGTTGATCATGATGCTGCTGCTGCTGCTGCTG 2759  
DB 963 CTATCTCAGGATCTGATCTTATTTGTTGATCATGATGCTGCTGCTGCTGCTGCTG 904  
QY 2760 TAACTAAGATTAAGGAGGCTTACATCTGAGCCGAGTGTGAAATGAAATGAAATG 2819  
DB 903 TAACTAAGATTAAGGAGGCTTACATCTGAGCCGAGTGTGAAATGAAATGAAATG 844  
QY 2820 CAGGCTACCGGCTCCAGATTTATCAGCAATTAACAGCAGCGGAGAGGCGAGCGA 2879  
DB 843 CAGGCTACCGGCTCCAGATTTATCAGCAATTAACAGCAGCGGAGAGGCGAGCGA 784  
QY 2880 GAAATGCTCTGCAACTTTATCCGCTCCATCCAGCTTAATTAATTTGGCGGAGGTA 2939  
DB 783 GAAATGCTCTGCAACTTTATCCGCTCCATCCAGCTTAATTAATTTGGCGGAGGTA 724  
QY 2940 GAGTAAGTATGCGCAAGTTTAATGTTTGGCAAGCTTTGCTCAATGCTCAAGGATTCG 2999  
DB 723 GAGTAAGTATGCGCAAGTTTAATGTTTGGCAAGCTTTGCTCAATGCTCAAGGATTCG 664  
QY 3000 TGGGTACAGCTGTGCTTGTGTTATGCTTATGCTTATGCTTATGCTTATGCTTATG 3059  
DB 663 TGGGTACAGCTGTGCTTGTGTTATGCTTATGCTTATGCTTATGCTTATGCTTATG 604  
QY 3060 GAGTTACATGATCCCGCATGTTGTGCAAAAAAGCGGTTAGCTCTTGTGCTCCGATCG 3119

Db 603 GAGTAACTGATCCCGCATGTTGTGCAAAAAACCGGTTAGCTCCCTCCGATCG 544.  
Qy 3120 TTGTGAGAAGTAAGTTGGCCGAGTGTATCACTAGGTTATGGCAGCATGCATTAAT 3179  
Db 543 TTGTGAGAAGTAAGTTGGCCGAGTGTATCACTAGGTTATGGCAGCATGCATTAAT 484  
Qy 3180 CTCTTACGTGATGTCATCCGTAAGATGCTTTCTGTGATCGGTAGTATCAACCAAGT 3239  
Db 483 CTCTTACGTGATGTCATCCGTAAGATGCTTTCTGTGATCGGTAGTATCAACCAAGT 424  
Qy 3240 CATTCTGAGAATAGTGTATGCGCGAGTGTCTTCTGCGCGGCTCAATACGGAGTA 3299  
Db 423 CATTCTGAGAATAGTGTATGCGCGAGTGTCTTCTGCGCGGCTCAATACGGAGTA 364  
Qy 3300 ATACCGCGCCACATAGCAGAACTTTAAAGTCTCATCATGGAACGTTCTCGGGGC 3359  
Db 363 ATACCGCGCCACATAGCAGAACTTTAAAGTCTCATCATGGAACGTTCTCGGGGC 304  
Qy 3360 GAAAACCTCAAGATCTTTACCGCTGTGAGATCCAGTTCCATGTTAACCCACTCGTGCAC 3419  
Db 303 GAAAACCTCAAGATCTTTACCGCTGTGAGATCCAGTTCCATGTTAACCCACTCGTGCAC 244  
Qy 3420 CCAACTGATCTTACAGCATCTTTTACTTTCACCAAGGCTTCTGGTGAGCAAAAAACAGGA 3479  
Db 243 CCAACTGATCTTACAGCATCTTTTACTTTCACCAAGGCTTCTGGTGAGCAAAAAACAGGA 184  
Qy 3480 GGCAGAAATGCGCAAAAAAGGAATAGGCGCGACACAGGAATGTTGAATACATACACT 3539  
Db 183 GGCAGAAATGCGCAAAAAAGGAATAGGCGCGACACAGGAATGTTGAATACATACACT 124  
Qy 3540 TCCCTTTTCAATATATTAAGACATTTTATCAGGTTATTTGTCTCATGACGGATACAT 3599  
Db 123 TCCCTTTTCAATATATTAAGACATTTTATCAGGTTATTTGTCTCATGACGGATACAT 64  
Qy 3600 TTGAATGATTTAGAAAAATAAACAATAGGGGTTCCGCGACATTTCCCGAAAAAGTGC 3659  
Db 63 TTGAATGATTTAGAAAAATAAACAATAGGGGTTCCGCGACATTTCCCGAAAAAGTGC 4  
Qy 3660 CAC 3662  
Db 3 CAC 1

RESULT 7  
US-10-066-390-4/c  
; Sequence 4, Application US/10066390  
; Publication No. US20020146732A1  
; GENERAL INFORMATION:  
; APPLICANT: Padgett, Hal S.  
; APPLICANT: Lindbo, John A.  
; APPLICANT: Fitzmaurice, Wayne P.  
; TITLE OF INVENTION: A Method of Increasing Complementarity  
; FILE REFERENCE: In A Heteroduplex  
; CURRENT APPLICATION NUMBER: US/10/066,390  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 3637  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Synthetic construct  
US-10-066-390-4

Query Match 72.8%; Score 2664.8; DB 13; Length 3637;  
Best Local Similarity 83.8%; Pred. No. 0;  
Matches 3070; Conservative 0; Mismatches 567; Indels 25; Gaps 4;

Qy 1 CTAATGTGAAGCTTAATATTTGTTAAATTCGGTTAAATTTTGTAAATCAGCTC 60  
|||||

Db 3637 CTAATGTGAAGCTTAATATTTGTTAAATTCGGTTAAATTTTGTAAATCAGCTC 3578  
Qy 61 ATTTTAAACCAATAGCCGAAATCGCAAAATCCCTTATATAATCAAAAGATAGACCGA 120  
Db 3577 ATTTTAAACCAATAGCCGAAATCGCAAAATCCCTTATATAATCAAAAGATAGACCGA 3518  
Qy 121 GATAGGGTTGAGTGTGTTCCAGTTGGAAACAAGTCCACTATTTAAAGAACGTGGACTC 180  
Db 3517 GATAGGGTTGAGTGTGTTCCAGTTGGAAACAAGTCCACTATTTAAAGAACGTGGACTC 3458  
Qy 181 CAAGTCAGAGGGGCAAAAAACGTCATACAGGGGATGCGCCACTAGCTGAACATCAC 240  
Db 3457 CAAGTCAGAGGGGCAAAAAACGTCATACAGGGGATGCGCCACTAGCTGAACATCAC 3398  
Qy 241 CTAATCAAGTTTTTGGGGTGTGAGGTGCGTAAAGCACTAAATCGAAACCTTAAGGGAG 300  
Db 3397 CTAATCAAGTTTTTGGGGTGTGAGGTGCGTAAAGCACTAAATCGAAACCTTAAGGGAG 3338  
Qy 301 CCCCCGATTTAGACCTTGAACGGGGGAAACGGCGCAACGTGGCGAAGAAAGGAAAGAA 360  
Db 3337 CCCCCGATTTAGACCTTGAACGGGGGAAACGGCGCAACGTGGCGAAGAAAGGAAAGAA 3278  
Qy 361 AGCGAAAGAGAGCGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCAAGCTGCGGTAAACAC 420  
Db 3277 AGCGAAAGAGAGCGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCAAGCTGCGGTAAACAC 3218  
Qy 421 CACACCGCGCGCTTAATGCGCGCTACAGAGGCGGTCCCATTCGCCATTCAAGCTGCG 480  
Db 3217 CACACCGCGCGCTTAATGCGCGCTACAGAGGCGGTCCCATTCGCCATTCAAGCTGCG 3158  
Qy 481 CAACCTGTGGGAAGGGGATGCGTGGCGGCGCTCTTCGCTATTAAGCCAGCTGGGAAAG 540  
Db 3157 CAACCTGTGGGAAGGGGATGCGTGGCGGCGCTCTTCGCTATTAAGCCAGCTGGGAAAG 3098  
Qy 541 GGGATGTGCTGCAAGCGATTAAGTGGGTAAACGCAAGGTTTTCCAGTCAAGCTGG 600  
Db 3097 GGGATGTGCTGCAAGCGATTAAGTGGGTAAACGCAAGGTTTTCCAGTCAAGCTGG 3038  
Qy 601 TAAACGACGCGCAGTGAACGCGCGCTGTCATTACAGTTTGAACCCGTGAGAGACG 660  
Db 3037 TAAACGACGCGCAGTGAACGCGCGCTGTCATTACAGTTTGAACCCGTGAGAGACG 2978  
Qy 661 GCACATCGCGGTGCAATGTGTTTACAGGTGATGAGAGATGAAGATGCTGCACAC 720  
Db 2977 CCGCGGTGCGG-----CCGCTGTGAACATGAGATCCG--CGG 2939  
Qy 721 GCTCAGAAACACGAGCTAGATTAACCTAGAAAGATATCATATTGAGAGTACGTTAA 780  
Db 2938 GCTCAGAAATCTTATTTGTATGTATGATTCATCCATGATGTATCCAGCAGAGTTA 2879  
Qy 781 AGATTAATCATGCGTAAATTTGACGATGGAATCTGTAAATGACTCATATAGGCGAAT 840  
Db 2878 CAAATCAAGAAAGACCATGTGATGATCGCTTTGTTGGATCTTTGAAAGGCAAGTT 2819  
Qy 841 TGGGTACCGGGCCCCCTCGAGGTGACAGGTATGCAATAGCTTGAATATGAAATCTCTGC 900  
Db 2818 GTGTGCAAGGTATGTTGTCTGTGTAAGAAAGAGGCGCATCCCAATGGAATATTTT 2759  
Qy 901 AGCCGGGGGATCACTAGTTCTTAGAGCGGCGCCACCGGGGAGAGCTCAGCTTTGT 960  
Db 2758 GTTATATATGTGTCTGATGTTGAACGATCATTTCAATGTGTGCGAATTTTGAAGT 2699  
Qy 961 TCCCTTATAGAGGTTAATTAATGATCCATGCGTCAATTTTACGACAGACTATCTTGTAG 1020  
Db 2698 TAGCTTATGATTCATCTTTGTTGTGCTGCGGATGATATCATGTGTGAG--TTATAG 2640  
Qy 1021 GGTAAATCTAGCTGATCAGATATATCGTGGGTCTTTTCCGGCTAGATCATCGCC 1080  
Db 2639 TTGTACTGTGAGTTGTGTGTCGAAATGTTCATCTTTTAAATCAATACCTTTTAAC 2580  
Qy 1081 CAAGCTGCGCTATCTGGGATCGGGAGGAGAAAGCCGCTGCTTTCCGCGAGGTTG 1140  
Db 2579 TCGATACGATTAACAGGGTATCACCTTCAAACTTACCTTCAGCAGCGGCTTGTAGTTC 2520



Qy 1141 AAGCGCATGAAAGATTGCGAGATGACTGCTGATTAAGCTTGAACGAAAC 1200  
Db 2519 CGGTACTTTTAAAGATATAGTGCTTCTGTATTAATACCTTTCGGCATGCGACTTTG 2460  
Qy 1201 GCACTTTTAAATGATATGATGAGGAGTGTGCGATGCAAGCTTTTAAACGATGATCTGT 2460  
Db 2459 AAAAGTATGACCGTTTATATGATCCGGATTAACGGGAAAAGCATTAACCATTAAG 2400  
Qy 1261 TCGTTACAGGCAACCGTGGATACCACTTTCGTGCGAGCTTTTTCGGAACAGAGTTCCGGATG 2320  
Db 2399 AAAAGTATGACCGTGGATACCACTTTCGTGCGAGCTTTTTCGGAACAGAGTTCCGGATG 2320  
Qy 1321 TCACCCGAAAGCGATCAGCAACCGAAACATTCGCGGACAGCCGGAATCGCGTCCG 1380  
Db 2339 AGGTTAAGCTTTCGTATGATGATCACTTTCACCTCTCCACTGACGAAATTTGTGC 2280  
Qy 1381 GTGTGAGATTAAATGACAGCGGTGCGCGCTGGATATTAAGTACGAGGAGGAGATC 1440  
Db 2279 CCAATTAACATCAACATTAATTAACAAGAAATGGGACATCTCAAGTGAATTTCTCT 2220  
Qy 1441 CTGGCTGATGCGGAGAAATGACATGATATACCCCGGATTAACCGCGGCGCGCTT 1500  
Db 2219 C---CTTTACTATGCTGATACCAAGCTTTTGTTCCTTTATGAGAGTTAATTCGCGCTT 2163  
Qy 1501 GCGTATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2163  
Db 2162 GCGTATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2103  
Qy 1561 CAACATAGAGCCCGAAGCATTAAGTAAAGCTGAGGAGCTTAATGATGATGATGATGATGAT 1620  
Db 2102 CAACATAGAGCCCGAAGCATTAAGTAAAGCTGAGGAGCTTAATGATGATGATGATGATGAT 2043  
Qy 1621 CACATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2043  
Db 2042 CACATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1880  
Qy 1681 GCATTAATGATGCGGCAACCGCGGAGAGCGGCTTTGCGTATGAGGCGCTTTCCG 1740  
Db 1982 GCATTAATGATGCGGCAACCGCGGAGAGCGGCTTTGCGTATGAGGCGCTTTCCG 1923  
Qy 1741 TTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800  
Db 1922 TTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1863  
Qy 1801 CTCAAAGCGGATTAACCGGTTATCCAGATCAGAGGATTAACGAGAAAGAACTATG 1860  
Db 1862 CTCAAAGCGGATTAACCGGTTATCCAGATCAGAGGATTAACGAGAAAGAACTATG 1803  
Qy 1861 AGCAAAAGCGGCAAAAGGCGAGAAACCGTAAAGGCGCGCTTCTGCTGCTTTTCA 1920  
Db 1802 AGCAAAAGCGGCAAAAGGCGAGAAACCGTAAAGGCGCGCTTCTGCTGCTTTTCA 1743  
Qy 1921 TAGGCTCGCCCTGAGCGATCAAAATTCAGCGCTCAAGTACAGAGGATGCGAA 1980  
Db 1742 TAGGCTCGCCCTGAGCGATCAAAATTCAGCGCTCAAGTACAGAGGATGCGAA 1683  
Qy 1981 CCGGACAGGACTTAAAGATTAACAGCGCTTCCCTGGAAGCTTCTGCTGCTGCTTCC 2040  
Db 1682 CCGGACAGGACTTAAAGATTAACAGCGCTTCCCTGGAAGCTTCTGCTGCTGCTTCC 1623  
Qy 2041 TGTTCGACCGCTGCGCTTAACGGATACCTGCTGCTTCTCTCTGCGGAAGCGTGC 2100  
Db 1622 TGTTCGACCGCTGCGCTTAACGGATACCTGCTGCTTCTCTCTGCGGAAGCGTGC 1563  
Qy 2101 GCTTTCTCATAGTCAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2160  
Db 1562 GCTTTCTCATAGTCAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1503  
Qy 2161 GGGCTGTGTCAGAAACCGCGTTACGCGGAGCGCTTATCCGCTTATCCGCTTATCC 2220  
Db 1502 GGGCTGTGTCAGAAACCGCGTTACGCGGAGCGCTTATCCGCTTATCCGCTTATCC 1443

Qy 2221 TCTTATGTCAAACCGGATTAAGCAAGCACTTATGCGCTGCGACAGCCACTGATTAACAG 2280  
Db 1442 TCTTATGTCAAACCGGATTAAGCAAGCACTTATGCGCTGCGACAGCCACTGATTAACAG 1383  
Qy 2281 GATTAGCAGAGGAGATTAAGGCGGCTGCTTACAGAGTTCTTGAATGAGGAGCTTACCTA 2340  
Db 1382 GATTAGCAGAGGAGATTAAGGCGGCTGCTTACAGAGTTCTTGAATGAGGAGCTTACCTA 1323  
Qy 2341 CGGCTACCTAAGAGCAGATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 2400  
Db 1322 CGGCTACCTAAGAGCAGATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 1263  
Qy 2401 AAAAGAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2460  
Db 1262 AAAAGAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1203  
Qy 2461 TGTTCGACAGCAGATTAACCGGCAAGAAACCAACCGCTGATGAGGATTTGCTGATCTT 2520  
Db 1202 TGTTCGACAGCAGATTAACCGGCAAGAAACCAACCGCTGATGAGGATTTGCTGATCTT 1143  
Qy 2521 TGTTCGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2580  
Db 1142 TGTTCGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1083  
Qy 2581 ATTATCAAAAGGATCTTACCTATGATCTTAAATTAATTAATTAATTAATTAATTAAT 2640  
Db 1082 ATTATCAAAAGGATCTTACCTATGATCTTAAATTAATTAATTAATTAATTAATTAAT 1023  
Qy 2641 CTAAAGATTAATGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2700  
Db 1022 CTAAAGATTAATGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 963  
Qy 2701 TATCTAGGCAATCTGCTATTTTGTTCATCAATGATGATGATGATGATGATGATGATGAT 2760  
Db 962 TATCTAGGCAATCTGCTATTTTGTTCATCAATGATGATGATGATGATGATGATGATGAT 903  
Qy 2761 AACTAGATTAAGGAGGCTTACCAATGCGCCCGAGCTGCAATGATTAACCGGAGAGCC 2820  
Db 902 AACTAGATTAAGGAGGCTTACCAATGCGCCCGAGCTGCAATGATTAACCGGAGAGCC 843  
Qy 2821 ACGCTACCGGCTCCAGATTTATGCAATTAACAGCAGCGGAGGAGCGAGCGAG 2880  
Db 842 ACGCTACCGGCTCCAGATTTATGCAATTAACAGCAGCGGAGGAGCGAGCGAGCGAG 783  
Qy 2881 AAGTGTCTGCACTTATCGGCTCATGAGTCAATTAATGATGATGATGATGATGATGAT 2940  
Db 782 AAGTGTCTGCACTTATCGGCTCATGAGTCAATTAATGATGATGATGATGATGATGAT 723  
Qy 2941 AGTAAGTATGCGCAAGTAAATGTTGCGCAAGCTTGTGCAATGCTACAGGAGCTAG 3000  
Db 722 AGTAAGTATGCGCAAGTAAATGTTGCGCAAGCTTGTGCAATGCTACAGGAGCTAG 663  
Qy 3001 GGTGTCAGGCTGCTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3060  
Db 662 GGTGTCAGGCTGCTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 603  
Qy 3061 AGTTACATGATCCCGCATGTTGTCAGAAAGCGGTTAGCTTCTGCTGCTGCTGCTGCT 3120  
Db 602 AGTTACATGATCCCGCATGTTGTCAGAAAGCGGTTAGCTTCTGCTGCTGCTGCTGCT 543  
Qy 3121 TGTTCAGAGTAAAGTTGCGCGAGTGTATCACTAGGTTATGCGACACTGCAATTAATC 3180  
Db 542 TGTTCAGAGTAAAGTTGCGCGAGTGTATCACTAGGTTATGCGACACTGCAATTAATC 483  
Qy 3181 TCTTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3240  
Db 482 TCTTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 423  
Qy 3241 ATTTCGAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3300  
Db 422 ATTTCGAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 363  
Qy 3301 TACGCGGCACATGACAGAACTTTAAAGTGCATCATTTGAGAAACGTTCTTCCGCGCG 3360



Db 362 TACGCGCCACATGACGAACTTTAAAGTCTCATCTTGGAAAACGTTCTTGGGGCG 303

Qy 3361 AAAACTCACAAGATCTTACCGCGTGTAGATTCAGTTCGATGTATCCCACTCGGCAAC 3420

Db 302 AAAACTCTCAAGATCTTACCGCGTGTAGATTCAGTTCGATGTATCCCACTCGGCAAC 243

Qy 3421 CACTGATCTTCAAGCATCTTTTACTTTCACACAGCGTTCTGGGTGAGCAAAAACGAGAG 3480

Db 242 CACTGATCTTCAAGCATCTTTTACTTTCACACAGCGTTCTGGGTGAGCAAAAACGAGAG 183

Qy 3481 GCAAAATGCGCCAAAAGGGAATPAGGGCGACACGAAATGTTGAATATCTCATACTCTT 3540

Db 182 GCAAAATGCGCCAAAAGGGAATPAGGGCGACACGAAATGTTGAATATCTCATACTCTT 123

Qy 3541 CCTTTTCAATATTATTGAGCATTTATCAGGGTATTGTCTCATGACGGATACATATT 3600

Db 122 CCTTTTCAATATTATTGAGCATTTATCAGGGTATTGTCTCATGACGGATACATATT 63

Qy 3601 TGAATGTTTATGAAAAATTAACAATAGGGGTTCCGGCGACATTTCCCGAAAAAGTGCC 3660

Db 62 TGAATGTTTATGAAAAATTAACAATAGGGGTTCCGGCGACATTTCCCGAAAAAGTGCC 3

Qy 3661 AC 3662

Db 2 AC 1

```

RESULT 8
US-10-206-030-4/C
; Sequence 4, Application US/10206030
; Publication No. US20020177160A1
; GENERAL INFORMATION:
; APPLICANT: Padgett, Hal S.
; APPLICANT: Lindbo, John A.
; APPLICANT: Fitzmaurice, Wayne P.
; TITLE OF INVENTION: A Method of Increasing Complementarity
; TITLE OF INVENTION: In A Heteroduplex
; FILE REFERENCE: P-LG 4878
; CURRENT APPLICATION NUMBER: US/10/206,030
; CURRENT FILING DATE: 2002-07-25
; PRIOR APPLICATION NUMBER: US/10/066,390
; PRIOR FILING DATE: 2002-02-01
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 3637
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic construct
US-10-206-030-4

```

Query Match	72.8%	Score 2664.8	DB 13	Length 3677
Best Local Similarity	83.8%	Pred. No. 0		
Matches 3070	Conservative 0	Mismatches 567	Indels 25	Gaps 4

Qy	1	CTAAATTGTAGCGTTAATATTTTGTGTTAAATTCGCGTTAAATTTTGTGTTAAATCAGCTC	60
Db	3637	CTAAATTGTAGCGTTAATATTTTGTGTTAAATTCGCGTTAAATTTTGTGTTAAATCAGCTC	3578
Qy	61	ATTTTAAACCAATAGGCGCGAATTCGGGAAATCCCTATTAATCAAAAGATGACCGA	120
Db	3577	ATTTTAAACCAATAGGCGCGAATTCGGGAAATCCCTATTAATCAAAAGATGACCGA	3518
Qy	121	GATAGGGTTGAGTGTGTTCCAGTTGGAAACAAGTCCACTATTAAGAACGTGACCTC	180
Db	3517	GATAGGGTTGAGTGTGTTCCAGTTGGAAACAAGTCCACTATTAAGAACGTGACCTC	3458
Qy	181	CAAGCTCAAAAGGGGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACATCAC	240
Db	3457	CAAGCTCAAAAGGGGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACATCAC	3398

QY	241	CTAATCAAGTTTTTGGGGTGAAGTGCCTGTAAGCACTTAATCGGAACCTTAAGGAG	300
Db	3397	CTAAACAAAGTTTTTGGGGTGAAGTGCCTTAAGCACTTAATCGGAACCTTAAGGAG	3336
QY	301	CCCCGATTTTAGAGCTTGAACGGGGAAACCGGGCGAAGTGGCGAAGAAAGGAAGAA	360
Db	3337	CCCCGATTTAGAGCTTGAACGGGGAAACCGGGCGAAGTGGCGAAGAAAGGAAGAA	3278
QY	361	AGCGAAGAGACGGGCGCTAGGGGCTGGCAAGTGTAGCGGTCAAGCTGACGCTGAACAC	420
Db	3277	AGCGAAGAGACGGGCGCTAGGGGCTGGCAAGTGTAGCGGTCAAGCTGACGCTGAACAC	3218
QY	421	CACACCGCGCGCTTAATGCGCGCTACAGGGCGCGTCCCATTTGCACTTACGGTCG	480
Db	3217	CACACCGCGCGCTTAATGCGCGCTACAGGGCGCGTCCCATTTGCACTTACGGTCG	3158
QY	481	CAACGTTGGGAAGGGCGATCGGTGCGGGCGCTTCCGTATTACGCGACGTGGGAAG	540
Db	3157	CAACGTTGGGAAGGGCGATCGGTGCGGGCGCTTCCGTATTACGCGACGTGGGAAG	3098
QY	541	GGGATGTCTCCAAAGCGCATTAAGTTGGGTAAACCCAGGGTTTTCCAGTACAGCTTG	600
Db	3097	GGGATGTCTCCAAAGCGCATTAAGTTGGGTAAACCCAGGGTTTTCCAGTACAGCTTG	3038
QY	601	TAAACGACGGCCAGTAGAGCGCGCTCGTTCATTCACGTTTTGAACCCGTGAGAGCG	660
Db	3037	TAAACGACGGCCAGTAGAGCGCGCTCGTTCATTCAGCTACTATAGGGCGAATTGGAGCTCA	2978
QY	661	GCACACTCGCGGTCCAAATGTGTTTTACAGGTGATGGAAGCAAGATGCTGCACAC	720
Db	2977	CGCGGTGCGCG-----CGCTCTGAAGACTAGTGAGATCCCC--CGG	2939
QY	721	GCTGCGAAGACGCGAGTAGATTAAACCTGAAGAATATATATTTGACGTACGTAA	780
Db	2938	GCTGCGAAGATTTCTTATTTGTATAGTTATCATCAGTCCCATGTGTATCCGACAGACTTA	2879
QY	781	AGATTAATCATGCTAAATTGACGCATGGAATCTGTATAAGACTCATATAGGCGAAT	840
Db	2878	CAAACTCAAGAGGACATGTGTGCACGCTTTTGTTGGAGTCTTTGGAAGGCAAGATT	2819
QY	841	TGGGTACGGGGCCCCCTCGAGGTCCAGGATATCGATTAAGCTGATATGCAATTCCTGC	900
Db	2818	GTGTGCAAGGTAAATGTTGTCTGTAAAGGAAGGCGCATCGCCAAATGGAGTATTTT	2759
QY	901	AGCCCGGGGATTCACCTAGTCTPAAGCGGGCGGCAACGGGTGGAGCTCAGGCTTTGT	960
Db	2758	GTTGATTAATGATCTGTAGTGAACCGATTCATTTCAATGTTGGGAATTTTGAAGT	2659
QY	961	TCCCTTAGTAGAGGATTAATTAAGATCCCATGCGTCAATTTACGCAACTATCTTCTAG	1020
Db	2698	TAGCTTGATTCATTTCTTTTGTGTGTGCGGATGATACATTTGTGTAG--TTATAG	2640
QY	1021	GATTATCTAGCTGCATCAGGATCATATCGTGGGCTTTTTCCGGCTCAGTATGCGC	1080
Db	2639	TTGTACTGAGTTGTGTGCGAGATGTTTTCATCTTTTAAATCAATACCTTTTAAC	2580
QY	1081	CAAGGTGGCGTATCTGGGCAATGGGGAGAAAGCCGTCCTTTTCCGCGAGGTG	1140
Db	2579	TCGATACATTAACAGGGATATCACTTCAAACTTGAACCTTGACAACGCGCTTTGTATTC	2520
QY	1141	AAGCGGATGAAAGATTTGCGAGAGATGACTGCTGCTGATTAAGCTTGAAGGAAAC	1200
Db	2519	CCGTATCTTTGAAGAATATAGTCCGTTCCGTATCATPAACCTTGGGGATGGCACTGTG	2460
QY	1201	GCACGTTTACATATGATTCGGGAAGGTGTGSCCATGACGCTTTAAGGTGAACCTGT	1260
Db	2459	AAAAAGTCATCCCTTTTCATATGATCCGATPAACGGGAAAGCAATTAACACCAATAGAG	2400
QY	1261	TGTTTCAAGGACCGGGATACAGATTGTCGGGCTTTTCCGGAACAAGTCCGAGTGG	1320
Db	2339	AAAGTATGACAAAGTTTGGCCATGAAACAGTATGTTTTTCAGATATGCAATTAATTTA	2340
QY	1321	TCACCGGAACGCATACGAACCCGAACATACCGCGACAGCCGGAATCGCGTGC	1380

```
Db 2339 AGGTAAGCTTCCGATGTAAGCATCACTTCACTCCCTCCACTCAAGAAAATTTGNC 2280
QY 1381 GTGTGACATTAATGACAGCGGTGCGCGCTGGGATATTAAGTCAAGGAGACGGATATC 1440
Db 2279 CCAATTAATCAACCATTAATTAACAGAAATTTGGAGAACCTCCAGTGAAGAAATTTCTTC 2220
QY 1441 CTGGCTGGATGCCGAGAAATGACATGSAATACCCCGTAGTTACCCGGCGCGGCTT 1500
Db 2219 C---CTTTTACTACCGTAACCAAGCTTTTGTTCCTTTAGTAGAGGTTAATTTCCGCTT 2163
QY 1501 GGGGTAAATCATAGTCTATAGTCTTTCTGTGTGAAATTTGTTATCCGCTCAATTTCCACA 1560
Db 2162 GGGGTAAATCATAGTCTATAGTCTTTCTGTGTGAAATTTGTTATCCGCTCAATTTCCACA 2103
QY 1561 CAACATACAGCGCGAGACATTAAGTGAAGCTTGGAGTGCCTTAATGAGTAACT 1620
Db 2102 CAACATACAGCGCGAGACATTAAGTGAAGCTTGGAGTGCCTTAATGAGTAACT 2043
QY 1621 CAATTAATGCGTTGGGCTCACTGCCGCTTTCAGTCGGGAACTGTGCGAGCT 1680
Db 2042 CACATTAATTCGCTTGCCTCAGCTCCGCTTCCAGTCGGGAACTGTGCGAGCT 1983
QY 1681 GCAATTAATGATCGGCCAACGCGCGGGAGAGGCGGTTTGGATATGCGCGCTTTCGCG 1740
Db 1982 GCAATTAATGATCGGCCAACGCGCGGGAGAGGCGGTTTGGATATGCGCGCTTTCGCG 1923
QY 1741 TTCTCTGCTCACTGACTGCTGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800
Db 1922 TTCTCTGCTCACTGACTGCTGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1863
QY 1801 CTCAAAAGGCGGTAATACGTTATTCACAGAAATCAGGGGATTAAGCGAAAGACATGTG 1860
Db 1862 CTCAAAAGGCGGTAATACGTTATTCACAGAAATCAGGGGATTAAGCGAAAGACATGTG 1803
QY 1861 AGCAAAAGGCCAGCAAAAGGCCAGAAACGTTAAAGCCGCTGTGCTGCGCTTTTCCA 1920
Db 1802 AGCAAAAGGCCAGCAAAAGGCCAGAAACGTTAAAGCCGCTGTGCTGCGCTTTTCCA 1743
QY 1921 TAGGCTCCGCCCTCTACAGACATCACAAAAATCGACGCTCAAGTCAAGAGTGGCAAA 1980
Db 1742 TAGGCTCCGCCCTCTACAGACATCACAAAAATCGACGCTCAAGTCAAGAGTGGCAAA 1683
QY 1981 CCGACAGGACTAATAAGATACAGAGCGTTTCCCGCTGGAAGCTCTCTGCTGCGCTTCC 2040
Db 1682 CCGACAGGACTAATAAGATACAGAGCGTTTCCCGCTGGAAGCTCTCTGCTGCGCTTCC 1623
QY 2041 TGTTCGACCTGCGCTTACCGGATACCTGTCCGCTTCTCCCTTCCGGAAGCGTGC 2100
Db 1622 TGTTCGACCTGCGCTTACCGGATACCTGTCCGCTTCTCCCTTCCGGAAGCGTGC 2163
QY 2101 GCTTCTCATAGTCAAGCTGTAGATATCTCAGTTCGAGTGGAGTGTGCTGCTCAAGCT 2160
Db 1562 GCTTCTCATAGTCAAGCTGTAGATATCTCAGTTCGAGTGGAGTGTGCTGCTCAAGCT 1503
QY 2161 GGGCTGTGTGCAAGAACCCCGCTTCAAGCCGAGCGCTGCGCTTATCCGGTAACATA 2220
Db 1502 GGGCTGTGTGCAAGAACCCCGCTTCAAGCCGAGCGCTGCGCTTATCCGGTAACATA 1443
QY 2221 TCTTGAATCCAAACCGGTAAAGACAGCTTATGCGCATGCGACAGACAGCTGTAAAG 2280
Db 1442 TCTTGAATCCAAACCGGTAAAGACAGCTTATGCGCATGCGACAGACAGCTGTAAAG 1383
QY 2281 GATTAGAGAGCGAGTATGATGAGCGGTGCTACAGAGTCTTGAAGTGTGCGCTTAACTA 2340
Db 1382 GATTAGAGAGCGAGTATGATGAGCGGTGCTACAGAGTCTTGAAGTGTGCGCTTAACTA 1323
QY 2341 CGGCTCACTAGAGAGAGATTTTGGATATGCGCTGCTGAGAACCGAGTTAACCTTGG 2400
Db 1322 CGGCTCACTAGAGAGAGATTTTGGATATGCGCTGCTGAGAACCGAGTTAACCTTGG 1263
QY 2401 AAAAAGAGTGGTCTTGAATCGGCAAAACACGCGGTACCGGTGCTTTT 2460
Db 1262 AAAAAGAGTGGTCTTGAATCGGCAAAACACGCGGTACCGGTGCTTTT 1203
QY 2461 TGTTCGAGACGACAGATTAAGCGCAAGAAAAGATCTCAAGAAATCTTTGATCTT 2520
Db 1202 TGTTCGAGACGACAGATTAAGCGCAAGAAAAGATCTCAAGAAATCTTTGATCTT 1143
QY 2521 TTTCAAGGGGTCTGACCGCTGAGTGAAGCAAAAATCAGCTTAAAGGATTTTGTATGAG 2580
Db 1142 TTTCAAGGGGTCTGACCGCTGAGTGAAGCAAAAATCAGCTTAAAGGATTTTGTATGAG 1083
QY 2581 ATTATCAAAAAGATCTTCACTAGATCTTTTAAATTTAAATGAATTTTAAATCAAT 2640
Db 1082 ATTATCAAAAAGATCTTCACTAGATCTTTTAAATTTAAATGAATTTTAAATCAAT 1023
QY 2641 CTAAAGTATATATAGTAAACTTGTGTCAGCTTACCAATGCTTAAATCAGTGAAGCACC 2700
Db 1022 CTAAAGTATATATAGTAAACTTGTGTCAGCTTACCAATGCTTAAATCAGTGAAGCACC 963
QY 2701 TATCTAGCGATCTGCTATTTTGGTCAATCCATAGTTCCTGACCTCCGCTGCTAGAT 2760
Db 962 TATCTAGCGATCTGCTATTTTGGTCAATCCATAGTTCCTGACCTCCGCTGCTAGAT 903
QY 2761 AACTAGATACGGAGGCGCTTACATCTGSCCCCACTGCTGCAATGATACCGAGACCC 2820
Db 902 AACTAGATACGGAGGCGCTTACATCTGSCCCCACTGCTGCAATGATACCGAGACCC 843
QY 2821 AGCTACCGGCTCAGATTTATACGAAATTAACAGCAAGCGAGGAGGCGGAGCCGAG 2880
Db 842 AGCTACCGGCTCAGATTTATACGAAATTAACAGCAAGCGAGGAGGCGGAGCCGAG 793
QY 2881 AAGTGTCTCGCAACTTATTCGCTCCATCCAGCTATTAATTTGTCGGGAAGCTAG 2940
Db 792 AAGTGTCTCGCAACTTATTCGCTCCATCCAGCTATTAATTTGTCGGGAAGCTAG 723
QY 2941 AGTAAGTATTCGCCAGTAAATAGTTTGGCAAGGTTGTCAGATTTGTCGGGAAGCTAG 3000
Db 722 AGTAAGTATTCGCCAGTAAATAGTTTGGCAAGGTTGTCAGATTTGTCGGGAAGCTAG 663
QY 3001 GGTGTACGCTGCTGCTTGTGATGCTTCAATTCAGCTCGGTTGCCAAGATCAAGCG 3060
Db 662 GGTGTACGCTGCTGCTTGTGATGCTTCAATTCAGCTCGGTTGCCAAGATCAAGCG 603
QY 3061 AGTTACATGATCCCAATGTTGTGCAAAAAGGCGTTACTCTCTTCCGCTCCGATGCT 3120
Db 602 AGTTACATGATCCCAATGTTGTGCAAAAAGGCGTTACTCTCTTCCGCTCCGATGCT 543
QY 3121 TGTCAAGATTAATGAGCGGAGCTGTTATCACTCAATGTTATGCAAGACAGCTAATTC 3180
Db 542 TGTCAAGATTAATGAGCGGAGCTGTTATCACTCAATGTTATGCAAGACAGCTAATTC 483
QY 3181 TCTTACGTGATGCCATCCGTAAAGATGCTTTTCTGTGATCGGTGATGATCAACCAAGTC 3240
Db 482 TCTTACGTGATGCCATCCGTAAAGATGCTTTTCTGTGATCGGTGATGATCAACCAAGTC 423
QY 3241 ATTCTGAATATGATATGCGGAGCGAGCGAGTGTCTTCCGCGGCTCAATACCGGATTA 3300
Db 422 ATTCTGAATATGATATGCGGAGCGAGCGAGTGTCTTCCGCGGCTCAATACCGGATTA 363
QY 3301 TACCGGCGCACTAGCAGAACTTTAAAGTCTCATATGGAATAAGTCTTCCGCGCG 3360
Db 362 TACCGGCGCACTAGCAGAACTTTAAAGTCTCATATGGAATAAGTCTTCCGCGCG 303
QY 3361 AAAAATCTCAAGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCATCTGTGAC 3420
Db 302 AAAAATCTCAAGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCATCTGTGAC 243
QY 3421 CAATGATCTTCAAGATCTTATCTTTCACAGCGTTTCTGGGTGACCAAAAAGAGAG 3480
Db 242 CAATGATCTTCAAGATCTTATCTTTCACAGCGTTTCTGGGTGACCAAAAAGAGAG 183
QY 3481 GCAAAAATGCGGCAAAAAGAGATTAAGGCGACACGGAATGTTGAATCTATATCTTT 3540
Db 182 GCAAAAATGCGGCAAAAAGAGATTAAGGCGACACGGAATGTTGAATCTATATCTTT 123
```

Query 3541 CCTTTTGAATATTTATTAAGCATTTATCAAGGTTATTTCTCATGAGCGATATATAT 3600  
Db 122 CCTTTTGAATATTTATTAAGCATTTATCAAGGTTATTTCTCATGAGCGATATATATAT 63  
Query 3601 TGAATGTATTTAGAAAATAAACAATAAGGGTCCGCGACATTTCCCGAAAGTGCC 3660  
Db 62 TGAATGTATTTAGAAAATAAACAATAAGGGTCCGCGACATTTCCCGAAAGTGCC 3  
Query 3661 AC 3662  
Db 2 AC 1  
RESULT 9  
US-10-211-079-4/c  
Sequence 4, Application US/10211079  
Publication No. US20030148315A1  
GENERAL INFORMATION:  
APPLICANT: Padgett, Hal S.  
APPLICANT: Vaewhongs, Andrew A.  
APPLICANT: Vojdani, Fakhrul S.  
APPLICANT: Smith, Mark L.  
TITLE OF INVENTION: Nucleic Acid Molecules Encoding CEL I  
TITLE OF INVENTION: Endonuclease and Methods of Use Thereof  
FILE REFERENCE: P-1G 5381  
CURRENT APPLICATION NUMBER: US/10/211,079  
PRIOR FILING DATE: 2002-08-01  
PRIOR APPLICATION NUMBER: US 10/098,155  
NUMBER OF SEQ ID NOS: 29  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 4  
LENGTH: 3637  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: synthetic construct  
US-10-211-079-4  
Query Match 72.8%; Score 2664.8; DB 15; Length 3637;  
Best Local Similarity 83.8%; Pred. No. 0;  
Matches 3070; Conservative 0; Mismatches 567; Indels 25; Gaps 4;  
Query 1 CTTAATTTGAACGCTTAATATTTGTTAAATTCGCGTTAAATTTTGTAAATCAGCTC 60  
Db 3637 CTTAATTTGAACGCTTAATATTTGTTAAATTCGCGTTAAATTTTGTAAATCAGCTC 3578  
Query 61 ATTTTAAACCAATAGGCGGAAATCGGAAATCCCTTAATCAAAAGAAATGACCGA 120  
Db 3577 ATTTTAAACCAATAGGCGGAAATCGGAAATCCCTTAATCAAAAGAAATGACCGA 3518  
Query 121 GATAGGGTGTAGTGTTCAGAGTTTGAACAAGATCCACTATTTAAAGACGTGACTC 180  
Db 3517 GATAGGGTGTAGTGTTCAGAGTTTGAACAAGATCCACTATTTAAAGACGTGACTC 3458  
Query 181 CAAGCTCAAGGCGGAAACCGCTCTATCAGGCGCATGCGCCACTAGCTGAACCATCAC 240  
Db 3457 CAAGCTCAAGGCGGAAACCGCTCTATCAGGCGCATGCGCCACTAGCTGAACCATCAC 3398  
Query 241 CTAATCAAGTTTGTGGGTGAGAGTGCGCTAAAGCACTAATTCGGAACCTTAAGGAG 300  
Db 3397 CTAATCAAGTTTGTGGGTGAGAGTGCGCTAAAGCACTAATTCGGAACCTTAAGGAG 3338  
Query 301 CCCCCGATTTAGAGCTTGAACGCGGAAACCGCGCAAGCTGCGAAGAAAGGAGAA 360  
Db 3337 CCCCCGATTTAGAGCTTGAACGCGGAAACCGCGCAAGCTGCGAAGAAAGGAGAA 3278  
Query 361 AGCGAAAGAGAGCGGCGCTAGGCGCTGCGCAAGTGTACGGTCAAGCTGCGGTAACAC 420  
Db 3277 AGCGAAAGAGAGCGGCGCTAGGCGCTGCGCAAGTGTACGGTCAAGCTGCGGTAACAC 3218  
Query 421 CACACCCGCGGCTTAATGCGCGCTACAGGCGCGTCCCATTCGCAATTCAGGCTGCG 480

Db 3217 CACACCCGCGGCTTAATGCGCGCTACAGGCGCGCTCCCATTCGCAATTCAGGCTGCG 3158  
Query 481 CAACTGTTGGAGAGCGGATGCGGCGCTCTTCTGCTATTAACGCGCTGCGGAAAG 540  
Db 3157 CAACTGTTGGAGAGCGGATGCGGCGCTCTTCTGCTATTAACGCGCTGCGGAAAG 3098  
Query 541 GGGATGTCTGCAAGGCGATTAAGTGGGTAAAGCCAGGGTTTCCAGTACAGACGTTG 600  
Db 3097 GGGATGTCTGCAAGGCGATTAAGTGGGTAAAGCCAGGGTTTCCAGTACAGACGTTG 3038  
Query 601 TAAACGACGCGCAGTAGCGCGCGCTGTTCACTACGTTTTTGAACCCGAGAGAGAG 660  
Db 3037 TAAACGACGCGCAGTAGCGCGCGCTGTTCACTACGTTTTTGAACCCGAGAGAG 2978  
Query 661 GCAGACTCGCGGTCAAAATGTGTTTAAACGCGTATGAGAGATGATGCTGACAC 720  
Db 2977 CCGCGGTGCGG-----CGCTCTAGAACTAGTGGATCCCC--CGG 2939  
Query 721 GCTGCAAGAACGCACTAGATTAACCTAGAAAGATTAATCATATTGTAGAGTACGTTA 780  
Db 2938 GCTGCAAGAAATCTTAATTTGTATGTATCATTCATGCGCATGTGTATCCAGACAGATTA 2879  
Query 781 AGATTAATCATGCTTAATAATGACGATGCGATCTGTATACGACTCACTAATAGGCGAAT 840  
Db 2878 CAACTCAAGAGAGACATGTGTCAGCTTTTCTTGGATCTTTGAAAGGCGAGATT 2819  
Query 841 TGGGTACCGGCGCCCGCTGAGGTGACGATGATGATGATGATGATGATGATGATGAT 900  
Db 2818 GTGTCAGAGGTATGTTGTCTGTGTAAGAGACAGGCGCATGCGCAATGAGATATTTT 2759  
Query 901 AGCCCGGGGATCACTAGTCTTAGAGCGCGCCAGCCGCGGTGAGTCCAGCTTTTGT 960  
Db 2758 GTTATATATGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2699  
Query 961 TCCCTTAGTAGAGGTTAATTAATGATCCATGCGTCAATTTTAAACGACATCTTTCTAG 1020  
Db 2698 TAGCTTTGATTTCACTTTCTTTGTTGTCGCGGATGATATCATTTGTTGATG--TTATAG 2640  
Query 1021 GGTATATCTAGCTGATCAAGATCATATGCTGCGGCTTTTTCGCGCTCATGATGCGC 1080  
Db 2639 TTGATCTGAGTTTGTGTCGAGATGATTTTCATCTTTTAAATCAATACCTTTTAAC 2580  
Query 1081 CAAGCTGCGCTATCTGCGCATCGGAGAGAGAAAGCCGCTTTTCCGCGAGGTG 1140  
Db 2579 TCGATAGATTAACAGAGGATCACTTCAAACTTCAACGACGCGCTTTGTAGTTC 2520  
Query 1141 AAGCGGATGGAAGAGTTTGGCGAGATGCTGCTGATGACGTTGAGCGGAAAC 1200  
Db 2519 CCGTATCTTTGAAGATATATGTCGTTCTGTATCATTAACCTTGGCGATGCACTCTTG 2460  
Query 1201 GCACGTTTACCATGATATTCGGAAGGTGCGCATGACGCTTTAACGCGTGAACGTG 1260  
Db 2459 AAAAAGTCAAGCGCTTATATGATTCGGAATACGGAAAGACATTAACCATTAAGAG 2400  
Query 1261 TCGTTACGCGCACTGGAATACAGTTCTGTCGCGCTTTTCCGGAACAGTTCCGAGTGG 1320  
Db 2399 AAAGTATGACAGATGTTGCGCATGGAACAGTATGTTTCCAGATGTCGAATTAATTTA 2340  
Query 1321 TCAGCCGGAAGCGCATCAGCAACCGAACAATACCGGCGACGCGGAACTGCGGCGCG 1380  
Db 2339 AGGTTAGCTTTTCCGTATGATGATCACTTCACTTCCCTCCTCAGACGAAATTTGTC 2280  
Query 1381 GTGTGAGATTAATGACAGCGGTGCGCGCTGAGATTTACGTTACGAGGAGAGCGGATATC 1440  
Db 2279 CCATTAATCATCAGCATTAATTAACAAGAAATGGGACAACTCAGAGGAAAGTTCTTCT 2220  
Query 1441 CTGCTGATGCGCGCAAAATGACATGATATCCCGTGATTTACCGGCGGCGGCGCTT 1500  
Db 2219 C---CTTACTCATGATGATCCAGCTTTTGTCTTCTTATGAGGTTAAATTTGCGCTT 2163  
Query 1501 GGCCTAATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1560



RESULT 10  
US-10-356-708-17/C  
Sequence 17, Application US/10356708  
Publication No. US20030157682A1  
GENERAL INFORMATION:  
APPLICANT: Large Scale Biology Corporation  
TITLE OF INVENTION: MISMATCH ENDONUCLEASES AND METHODS OF USE THEREOF  
FILE REFERENCE: P-IG 10100  
CURRENT APPLICATION NUMBER: US/10/356,708  
PRIOR FILING DATE: 2003-01-31  
PRIOR APPLICATION NUMBER: 60/353,722  
PRIOR FILING DATE: 2000-02-01  
PRIOR APPLICATION NUMBER: 10/098,155  
PRIOR FILING DATE: 2002-03-14  
PRIOR APPLICATION NUMBER: 10/211,079  
NUMBER OF SEQ ID NOS: 35  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 17  
LENGTH: 3637  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Encodes cycle 3 GFP  
US-10-356-708-17.

Query Match 72.8%, Score 2664.8, DB 15, Length 3637,  
Best Local Similarity 83.8%, Pred. No. 0,  
Matches 3070, Conservative 0, Mismatches 567, Indels 25, Gaps 4;

1 CTAATGTAGACGCTTAATTTGTTAAATTCGCTTAATTTGTTAAATTCAGTC 60  
3637 CTAATGTAGACGCTTAATTTGTTAAATTCGCTTAATTTGTTAAATTCAGTC 3578  
61 ATTTTAAACCAATAGGCGCAATTCGCAAAATCCCTTAATCAAAAGATAGACGA 120  
3577 ATTTTAAACCAATAGGCGCAATTCGCAAAATCCCTTAATCAAAAGATAGACGA 3518  
121 GATAGGGTGAAGTGTTCCTGATTTGGAACAAGATCCACTATTAAAGAGTGGACTC 180  
3517 GATAGGGTGAAGTGTTCCTGATTTGGAACAAGATCCACTATTAAAGAGTGGACTC 3458  
181 CAAGTCAAAAGGGGAAACCCGTCTATCAGGGCGATGCCCACTAGCTGAACCATCACC 240  
3457 CAAGTCAAAAGGGGAAACCCGTCTATCAGGGCGATGCCCACTAGCTGAACCATCACC 3398  
241 CTAATCAAGTTTGGGGTGAAGTGTTCCTGATTTGGAACAAGATCCACTATTAAAGAGTGGACTC 300  
3397 CTAATCAAGTTTGGGGTGAAGTGTTCCTGATTTGGAACAAGATCCACTATTAAAGAGTGGACTC 3338  
301 CCCCCGATTTAGACCTTGAACGGGGAACCCGCGCAACCTGCGCAAGAAAGAGGAAAGAA 360  
3337 CCCCCGATTTAGACCTTGAACGGGGAACCCGCGCAACCTGCGCAAGAAAGAGGAAAGAA 3278  
361 AGCGAAAGAGAGGGCGCTGAGGGCGCTGAGCAAGTGTAGCGGTCAACGCTGCGCTGAACAC 420  
3277 AGCGAAAGAGAGGGCGCTGAGGGCGCTGAGCAAGTGTAGCGGTCAACGCTGCGCTGAACAC 3218  
421 CACACCGCGCGCTTAATGCGCGCTTACAGGGCGCGTCCCATTCGCGCATTCAGGCTGCG 480  
3217 CACACCGCGCGCTTAATGCGCGCTTACAGGGCGCGTCCCATTCGCGCATTCAGGCTGCG 3158  
481 CAACGTGTGGAGAGGGCGATGATGCGGGCTCTTCGCTAATTAACGCGAGCTGGCGAAGG 540  
3157 CAACGTGTGGAGAGGGCGATGATGCGGGCTCTTCGCTAATTAACGCGAGCTGGCGAAGG 3098  
541 GCGATGTCTGCAAGCGCATTAAGTGTGATTAACGCGAGGCTTTTCCAGTCAACGAGCTTG 600  
3097 GCGATGTCTGCAAGCGCATTAAGTGTGATTAACGCGAGGCTTTTCCAGTCAACGAGCTTG 3038  
601 TAAACGACGCGCAGTGAAGCGCGCTGCTTCAATTCAGCTTTTGAACCGGTGAGAGACGG 660

3037 TAAACGACGCGCAGTGAAGCGCGCTGCTTCAATTCAGCTTTTGAACCGGTGAGAGACGG 2978  
661 GCAAGCTCGCGTGCAGCAATGCTTTTACAGCTGATGAGCAATGAAGATGCTGACAC 720  
2977 CCGCGGTGCGG-----CCCTCTAGAACTAGTGAATCCCTCC--CGG 2939  
721 GCTCAGAACGCGAGCTAGATTAACCTAGAAAGATTAATATATTGAGTGAAGTGA 780  
2938 GCTCAGCAATTTCTTAATTTGATAGTTTATCCATGCAAGCTGATATCCACAGAGTTA 2879  
781 AGATTAATCATGCTTAATTAATGACCATGAGTCTGTATACGACTCACTATAGGCGCAT 840  
2878 CAATCTAAGAAAGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2819  
841 TGGGTACCGGGCCCCCTGAGGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 900  
2818 GTGTGACAGGTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2759  
901 AGCCCGGGGATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960  
2758 GTTGAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2699  
961 TCCCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020  
2698 TAGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2640  
1021 GATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080  
2639 TTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2580  
1081 CAAGCTGCGCTATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140  
2579 TCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2520  
1141 AAGCGCATGAGAAAGATTTGCGCGAGATGATGATGATGATGATGATGATGATGATGATGAT 1200  
2519 CCGTATCTTTGAAGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2460  
1201 GCACGTTTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1260  
2459 AAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2400  
1261 TCGTACAGGCACTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320  
2399 AAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2340  
1321 TCAAGCGGAGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1380  
2339 AGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2280  
1381 GTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1440  
2279 CCAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2220  
1441 CTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1500  
2219 C--CTTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2163  
1501 GCGCTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1560  
2162 GCGCTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2103  
1561 CAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1620  
2102 CAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2043  
1621 CAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1680  
2042 CAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1983  
1681 GCATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1740  
1982 GCATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1923







CURRENT FILING DATE: 2002-07-25  
PRIOR APPLICATION NUMBER: US/10/066,390  
PRIOR FILING DATE: 2002-02-01  
NUMBER OF SEQ ID NOS: 15  
SOFTWARE: FASTSEQ for Windows Version 4.0  
SEQ ID NO 4  
LENGTH: 3637  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: synthetic construct  
US-10-205-772-4

Query Match 72.8%; Score 2664.8; DB 15; Length 3637;

Best Local Similarity 83.8%; Pred. No. 0; Mismatches 567; Indels 25; Gaps 4;

Matches 3070; Conservative 0; Pseudomatches 567; Indels 25; Gaps 4;

QY 1 CTAATTGTAGAGCGTTAATTTTGTAAATTCGCGTTAAATTTTGTAAATCAGCTC 60  
DB 3637 CTAATTGTAGAGCGTTAATTTTGTAAATTCGCGTTAAATTTTGTAAATCAGCTC 3578

QY 61 ATTTTAAACCAATAGCGCAAAATCGCAAAATCCCTTAAATCAAAAGAAATAGACGA 120  
DB 3577 ATTTTAAACCAATAGCGCAAAATCGCAAAATCCCTTAAATCAAAAGAAATAGACGA 3518

QY 121 GATAGGGTGTAGTGTGTTCAGATTGGAACAAGTCACTATTAAAGAACTGAGCTC 180  
DB 3517 GATAGGGTGTAGTGTGTTCAGATTGGAACAAGTCACTATTAAAGAACTGAGCTC 3458

QY 181 CAAGCTCAAAAGGCGAAAAACCGTCTATCAGAGCGATGCGCCACTAGCGTAACCATCAC 240  
DB 3457 CAAGCTCAAAAGGCGAAAAACCGTCTATCAGAGCGATGCGCCACTAGCGTAACCATCAC 3358

QY 241 CTAATCAAGTTTTTGGGGTGTAGGTGCGGTAAAGCACTAAATCGGAACCTTAAAGGAG 300  
DB 3397 CTAATCAAGTTTTTGGGGTGTAGGTGCGGTAAAGCACTAAATCGGAACCTTAAAGGAG 3338

QY 301 CCCCCGATTAGACTTGAACGGGGAAGACCGCCGAACGTGCGGAAGAAAGGAGAA 360  
DB 3337 CCCCCGATTAGACTTGAACGGGGAAGACCGCCGAACGTGCGGAAGAAAGGAGAA 3278

QY 361 AGCGAAAGAGAGCGGCGCTTAGGCGCTGCAAGTGTAGCGGTCAAGCTGCGGTAAACAC 420  
DB 3277 AGCGAAAGAGAGCGGCGCTTAGGCGCTGCAAGTGTAGCGGTCAAGCTGCGGTAAACAC 3218

QY 421 CACACCCGCGCGCTTAAATGCGCGCTACAGGCGCGCTCCATTGCGCAATTCAGGCTGCG 480  
DB 3217 CACACCCGCGCGCTTAAATGCGCGCTACAGGCGCGCTCCATTGCGCAATTCAGGCTGCG 3158

QY 481 CAATCTGTGGGAAGGCGATGCGGTGCGGCGCTTTCGCTATTAGCGCACTGCGGAAGG 540  
DB 3157 CAATCTGTGGGAAGGCGATGCGGTGCGGCGCTTTCGCTATTAGCGCGCACTGCGGAAGG 3098

QY 541 GGGATGTGCTCAAGAGCGATTAAATGAGGTAAAGCGAGGTTTTCCAGTCAAGCGTTG 600  
DB 3097 GGGATGTGCTCAAGAGCGATTAAATGAGGTAAAGCGAGGTTTTCCAGTCAAGCGTTG 3038

QY 601 TAAACAGACGCGCAGTGAAGCGCGCTGCTTCAATTCAGTTTTTGAACCCGCTGAGAGAG 660  
DB 3037 TAAACAGACGCGCAGTGAAGCGCGCTGCTTCAATTCAGTTTTTGAAGCGCTGAGAGAG 2978

QY 661 GCAGACTCGCGGTGCAAAATGTGTTTACAGCGTATGAGCAAGTGAAGTGTGAGAC 720  
DB 2977 CGCGGTGCGGTGCAAAATGTGTTTACAGCGTATGAGCAAGTGAAGTGTGAGAC 2939

QY 721 GCTCAGAAACAGCAGCTAGATTAACTTGAAGATTAATCATTTTGTGAGCTAGCTTAA 780  
DB 2938 GCTCAGAAACAGCAGCTAGATTAACTTGTATGATTCATCCATGCTGAGCTAGCTTAA 2879

QY 781 AGATTAATCAGCTTAAATTTGAGCAAGTGAATCTGTAAATAGACTCACTAATAGGCGAAT 840  
DB 2878 CAAATCAAGAAAGACCATGTGTACAGCTTTTGTGAGATCTTTGAAAGGCGAAGTT 2819

QY 841 TGGGTACCGGCGCCCCCTCGAGGTGACAGGTATGATTAAGCTTATGATATTCCTGCG 900  
DB 2818 GTGTGACAGGTAAATGTTGTCTGTAAAGGAGAGGCGCCATCGCCAAATGGAATTTT 2759

QY 901 AGCCCGGGGATCCACTAGTTCTAGAGCGGCGCCACCGCGGTGAGCTCCAGCTTTTGT 960  
DB 2758 GTTATTAATGCTCTGCTAGTTGAACGAGATTCATCTTCAATCTTGTGCGAATTTGAAGT 2699

QY 961 TCCCTTTAGTAGGGTTAAATAGATCCAGTGCATATTTTACAGCACTATCTTTCTAG 1020  
DB 2698 TAGCTTTAGTTCCATCTTTTGTGTTGTCTGCGGAGTGAATACATGTGTAG--TTATAG 2640

QY 1021 GGTAACTAGCTGATCAGATATCATATCGTGGGTCTTTTTCGGCTCATGCTATCGCC 1080  
DB 2639 TTGACTGAGATTTGTGTCCAGAAATGTTCATCTTTTAAATCAATACCTTTTAAAC 2580

QY 1081 CAACTGCGCGCTATCTGCGCATTCGAGGAGAGAAAGACCGGCTTTTCCCGCAGGTTG 1140  
DB 2579 TCGATAGGATTAACAAGGATATCATCTTCAAACTGACTGACGACCGCTTGTAGTTG 2520

QY 1141 AAGCGCATGGAAGAAGTTTCCGAGATGACTGCTGCTGATTAAGCTTGAAGCGAAAC 1200  
DB 2519 CCGTCACTTTGAAGAATATAGTGTCTCTGTATCATTAACCTTGGCGCATGCACTGTG 2460

QY 1201 GCAGCTTACCATGATATTCGGAAGTGTGCGCATGCAAGCCTTTAAACGCTGAACCTGT 1260  
DB 2459 AAAAAGTATGCGCTTCAATATGATTCGGAATACGAGAAAGCAATTAACCATTAAGAG 2400

QY 1261 TCGTTACAGGCGCACTGAGGATACAGATTGTCGCGCTTTTCCGAGACAGTTCCGATG 1320  
DB 2399 AAAGTATGATGCAAGTGTGCGCATGGAACAGTATGTTTTCAGATAGTGCATAATTTTA 2340

QY 1321 TCAGCCCGAAGCGCATGAGCAACCGGAACAATACGCGCGACAGCGGAACTGCGTCCG 1380  
DB 2339 AGGCTAGACTTTCCGTATGATGATCATCTTCAACCTCTCACTGACAGAAATTTGTGC 2280

QY 1381 GTGTGCAATTAATGACAGCGGTGCGCGCTGCGGATTTACGTACAGAGAGACGGATATC 1440  
DB 2279 CCAATTAATCATCACCATTAAATTCACAAAGATTTGGAACCACTCCAGTAAAGTTCTCT 2220

QY 1441 CTGCGTGGATGCGCAAAATGACATGATATACCCCTGAGTTTACCGGCGGCGCGCTT 1500  
DB 2219 C---CTTTATCTCATCGGTATCCAGCTTTTGTGCTCTTTAGTGAAGGTTAATGCGCTT 2163

QY 1501 GCGGTAAATCATGTGATAGCTGTTTCTGTGTGAATTTGTTATTCGCTCAAAATTCACA 1560  
DB 2162 GCGGTAAATCATGTGATAGCTGTTTCTGTGTGAATTTGTTATTCGCTCAAAATTCACA 2103

QY 1561 CAACATACAGACCGGAAGCATTAAGTAAAGCTGCGGTGCTTAAATGATGAGCTTAAT 1620  
DB 2102 CAACATACAGACCGGAAGCATTAAGTAAAGCTGCGGTGCTTAAATGATGAGCTTAAT 2043

QY 1621 CACATTAATGCGGTGCTGCTCACTGCGCTTTCAGTCCGGAACCTGTCGTGCTCAAGCT 1680  
DB 2042 CACATTAATGCGGTGCTGCTCACTGCGCTTTCAGTCCGGAACCTGTCGTGCTCAAGCT 1963

QY 1681 GCATTAATGATGCGGCAACGCGCGGAGAGAGCGGTTTCGTAATGAGCGCTTTCGCG 1740  
DB 1982 GCATTAATGATGCGGCAACGCGCGGAGAGAGCGGTTTCGTAATGAGCGCTTTCGCG 1923

QY 1741 TTCTGCTGCTAGTACCTGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800  
DB 1922 TTCTGCTGCTAGTACCTGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1863

QY 1801 CTCAAAAGCGGTAAATAGTTTATTCACAGAAATCAGGGATTAACGAGAAAGAAACATGTG 1860  
DB 1862 CTCAAAAGCGGTAAATAGTTTATTCACAGAAATCAGGGATTAACGAGAAAGAAACATGTG 1803

QY 1861 AGCAAAAGGCGCAAGAAAGGCGCAAGACCGTAAAGAGCGCGTGTGCTGCTTTTCA 1920  
DB 1802 AGCAAAAGGCGCAAGAAAGGCGCAAGACCGTAAAGAGCGCGTGTGCTGCTTTTCA 1743

QY 1921 TAGGCTCGGCGCCCTGACAGCATCAAAAATTCAGCTCAAGTCAAGTGTGCGGAA 1980

```

Db      1742 TAGGCTCCGCCCCCTGACGAGCATCAAAAATCAAGCTCAAGGTCAAGGTGGGAAA 1683
Oy      1981 CCCGACGAGACTATPAAAGATACACAGCGTTTCCCGCTGGAAGTCCCTGTGCTCC 2040
Db      1682 CCCGACGAGACTATPAAAGATACACAGCGTTTCCCGCTGGAAGTCCCTGTGCTCC 1623
Oy      2041 TGTTCGAGCCCTGCGCTTACCGGATACCTGTCCGCTTTCCTCCCTGGGAGCGTGC 2100
Db      1622 TGTTCGAGCCCTGCGCTTACCGGATACCTGTCCGCTTTCCTCCCTGGGAGCGTGC 1563
Oy      2101 GCTTCTCAGTCTCAAGCTGTAGATCTCAGTTCGAGTGTAGGTCGTTCCTCAAGCT 2160
Db      1562 GCTTCTCAGTCTCAAGCTGTAGATCTCAGTTCGAGTGTAGGTCGTTCCTCAAGCT 1503
Oy      2161 GGGCTGTGTGACGAAACCCCGCTTCAAGCGGACCGCTGCGCTTATCCGGTAACATCG 2220
Db      1502 GGGCTGTGTGACGAAACCCCGCTTCAAGCGGACCGCTGCGCTTATCCGGTAACATCG 1443
Oy      2221 TCTTGAATCCAAACCGGTAAACACAGTATCCGCTGAGAGCCACTGTGTAACG 2280
Db      1442 TCTTGAATCCAAACCGGTAAACACAGTATCCGCTGAGAGCCACTGTGTAACG 1383
Oy      2281 GATTAGCAGAGCAGAGTATGTAGGCGGTCTACAGAGTCTTGAAGTGTGCTTACTA 2340
Db      1382 GATTAGCAGAGCAGAGTATGTAGGCGGTCTACAGAGTCTTGAAGTGTGCTTACTA 1323
Oy      2341 CGGCTACCTAGAGAGACAGTATTTGTATCTGCTGCTGAGAGCCGTTACCTTGG 2400
Db      1322 CGGCTACCTAGAGAGACAGTATTTGTATCTGCTGCTGAGAGCCGTTACCTTGG 1263
Oy      2401 AAAAAGTGTGTAGCTCTTGATCCGCAAAACAAACACCGCTGTGTAGCGTGTGTTT 2460
Db      1262 AAAAAGTGTGTAGCTCTTGATCCGCAAAACAAACACCGCTGTGTAGCGTGTGTTT 1203
Oy      2461 TGTTCGAGCAGCAGATTAACGCGCAGAAAAGAGATCTCAAGAGATCTTGTGATCTT 2520
Db      1202 TGTTCGAGCAGCAGATTAACGCGCAGAAAAGAGATCTCAAGAGATCTTGTGATCTT 1143
Oy      2521 TTCTACGCGGCTGACGCTCAGTGGAGAAAACCTCAGCTTAAAGGATTTTGTGATGAG 2580
Db      1142 TTCTACGCGGCTGACGCTCAGTGGAGAAAACCTCAGCTTAAAGGATTTTGTGATGAG 1083
Oy      2581 ATTATCAAAAAGATCTTCACTAGATCTTTTAAATTTAAATGAAGTTTAAATCAAT 2640
Db      1082 ATTATCAAAAAGATCTTCACTAGATCTTTTAAATTTAAATGAAGTTTAAATCAAT 1023
Oy      2641 CTAAAGTATATATGAATTAACCTGATCTGACAGTTACATCTTATCAGTAGGACCC 2700
Db      1022 CTAAAGTATATATGAATTAACCTGATCTGACAGTTACATCTTATCAGTAGGACCC 963
Oy      2701 TATCTCAGCAGATCTGTCTATTTGCTTATCATATGTTGCTGCTGCTGCTGCTGCT 2760
Db      962 TATCTCAGCAGATCTGTCTATTTGCTTATCATATGTTGCTGCTGCTGCTGCTGCT 903
Oy      2761 AACTACGATACGGAGGCTTACCATCTGAGCCCACTGTGCAATGATACCGGAGACCC 2820
Db      902 AACTACGATACGGAGGCTTACCATCTGAGCCCACTGTGCAATGATACCGGAGACCC 843
Oy      2821 ACGCTCAGCGGCTCAGATTTATCAGCAATTAACCAAGCAGCGGAGGCGGAGCCAG 2880
Db      842 ACGCTCAGCGGCTCAGATTTATCAGCAATTAACCAAGCAGCGGAGGCGGAGCCAG 783
Oy      2881 AATGTGCTCTGCAACTTTATCCGCTTCACTCAGCTATTAATGTTGCGGAGAGCTAG 2940
Db      782 AATGTGCTCTGCAACTTTATCCGCTTCACTCAGCTATTAATGTTGCGGAGAGCTAG 723
Oy      2941 AGTAAGTATGTCGCAAGTTAAAGTTGGCAGAGCTGTGCTCAGAGCAATCGT 3000
Db      722 AGTAAGTATGTCGCAAGTTAAAGTTGGCAGAGCTGTGCTCAGAGCAATCGT 663
Oy      3001 GGTGTACGCTGTGTGTGTATGCTTCAATAGTCCGTTCCCAAGATCAAGGCG 3060

```

---

```

Db      662 GGTGTACGCTGTGTGTGTGTATGCTTATTCAGCTCCGATTCCCAAGATCAAGGCG 603
Oy      3061 AGTACATGATCCCGCATGTGTGCAAAAAGGGTATAGTCTTGTGCTTCCGATCGT 3120
Db      602 AGTACATGATCCCGCATGTGTGCAAAAAGGGTATAGTCTTGTGCTTCCGATCGT 543
Oy      3121 TGTCAAGATGATGTGGCCCGCATGTTTATCAGTATGATATGAGCAGACATGATTC 3180
Db      542 TGTCAAGATGATGTGGCCCGCATGTTTATCAGTATGATGATGAGCAGACATGATTC 483
Oy      3181 TCTTACTGTATGATCCATCCGTAAGATGCTTTCTGTGAGTGTGATCTCAACCAATC 3240
Db      482 TCTTACTGTATGATCCATCCGTAAGATGCTTTCTGTGAGTGTGATCTCAACCAATC 423
Oy      3241 ATTGTAGATGATGTATGATGCGGACCGAGTGTCTTGTCCCGGCTCAATAGGAGTAA 3300
Db      422 ATTGTAGATGATGTATGATGCGGACCGAGTGTCTTGTCCCGGCTCAATAGGAGTAA 363
Oy      3301 TACCGGCGCATATGAGCAATTTAAAGTGTCTATCAATGAAACAGTTCTTCCGAGCG 3360
Db      362 TACCGGCGCATATGAGCAATTTAAAGTGTCTATCAATGAAACAGTTCTTCCGAGCG 303
Oy      3361 AAAACTCTCAAGATCTTACCGCTGTGTGAGTCCAGTGTATGAAACCACTGTGACCC 3420
Db      302 AAAACTCTCAAGATCTTACCGCTGTGTGAGTCCAGTGTATGAAACCACTGTGACCC 243
Oy      3421 CAAGTATCTTCAAGATCTTCTTACCTTACCAAGGCTTCTGGGTGAGCAAAACAGGAG 3480
Db      242 CAAGTATCTTCAAGATCTTCTTACCTTACCAAGGCTTCTGGGTGAGCAAAACAGGAG 183
Oy      3481 GCAAAATGCGCAAAAAGGAAATAGGGGAGACAGGAAATGTTGATATCTACTCTCTT 3540
Db      182 GCAAAATGCGCAAAAAGGAAATAGGGGAGACAGGAAATGTTGATATCTACTCTCTT 123
Oy      3541 CTTTTTCAATATATTAAGAGATTTATCAGGTTATGTCATGAGCGGATACATAT 3600
Db      122 CTTTTTCAATATATTAAGAGATTTATCAGGTTATGTCATGAGCGGATACATAT 63
Oy      3601 TGAATGATTTAGAAAATTAACAAATAGGGGTTCCGCGCAATTTCCCGAAAAGTGC 3660
Db      62 TGAATGATTTAGAAAATTAACAAATAGGGGTTCCGCGCAATTTCCCGAAAAGTGC 3
Oy      3661 AC 3662
Db      2 AC 1

```

---

```

RESULT 12
US-10-280-913A-17/c
/ Sequence 17, Application US/10280913A
/ Publication No. US20040110130A1
/ GENERAL INFORMATION:
/ APPLICANT: Large Scale Biology Corporation
/ TITLE OF INVENTION: POPULATION OF POLYNUCLEOTIDE SEQUENCE VARIANTS
/ FILE REFERENCE: LSBC-PLG5392-CIP
/ CURRENT APPLICATION NUMBER: US/10/280,913A
/ PRIOR FILING DATE: 2002-10-25
/ PRIOR APPLICATION NUMBER: 60/402,342
/ PRIOR FILING DATE: 2002-08-08
/ PRIOR APPLICATION NUMBER: 10/066,390
/ PRIOR FILING DATE: 2002-02-01
/ PRIOR FILING DATE: 2001-02-14
/ PRIOR APPLICATION NUMBER: 60/266,386
/ PRIOR FILING DATE: 2002-02-02
/ NUMBER OF SEQ ID NOS: 35
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 17
/ LENGTH: 3637
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Encodes cycle 3 GFP

```

US-10-280-913A-17

Query Match 72.8%; Score 2664.8; DB 17; Length 3637;  
Bee Local Similarity 83.8%; Pred. No. 0;  
Matches 3070; Conservative 0; Mismatches 567; Indels 25; Gaps 4;

1 CTAATTTAGAGCGTTAATATTTTGTAAATTCGCGTTAAATTTTGTAAATCAGCTC 60  
3637 CTAATTTAGAGCGTTAATATTTTGTAAATTCGCGTTAAATTTTGTAAATCAGCTC 3578  
61 ATTTTAAACCAATAGCGCGAAATTCGCGAAATCCCTTAATTAATCAAAAGATAGACGA 120  
3577 ATTTTAAACCAATAGCGCGAAATTCGCGAAATCCCTTAATTAATCAAAAGATAGACGA 3518  
121 GATAGGGTTGAGTGTGTTCCAGTTTGAACAAGATCCACTAATTAAGAACTGTGACTC 180  
3517 GATAGGGTTGAGTGTGTTCCAGTTTGAACAAGATCCACTAATTAAGAACTGTGACTC 3458  
181 CAAGCTCAAGAGGCGGAAAAACGCTATCAAGGCGGATGCGCCACTAGCGTGAACCATCACC 240  
3457 CAAGCTCAAGAGGCGGAAAAACGCTATCAAGGCGGATGCGCCACTAGCGTGAACCATCACC 3398  
241 CTAATCAAGTTTGTGGGTCGAGGTGCGCTAAGCACTAATCGGAACCTTAAGGAGAG 300  
3397 CTAATCAAGTTTGTGGGTCGAGGTGCGCTAAGCACTAATCGGAACCTTAAGGAGAG 3338  
301 CCCCCGATTGAGCTTGAACGCGGGAAGCCGCGCAACGCTGCGGAGAAAGAGAGAA 360  
3337 CCCCCGATTGAGCTTGAACGCGGGAAGCCGCGCAACGCTGCGGAGAAAGAGAGAA 3278  
361 AGCGAAAGAGCGGCGCTAGGAGCGCTGCAAGGTATAGCGGTACGCGCGCGCTAACAC 420  
3277 AGCGAAAGAGCGGCGCTAGGAGCGCTGCAAGGTATAGCGGTACGCGCGCGCTAACAC 3218  
421 CACACCGCGCGCTTAAATGCGCGCTACAGGCGCGCTCCCATTCGCCATTCAAGCTCG 480  
3217 CACACCGCGCGCTTAAATGCGCGCTACAGGCGCGCTCCCATTCGCCATTCAAGCTCG 3158  
481 CAACCTGTGGGAAGGCGATGCGTGCAGGCGCTTCTTCCCTAATTAAGCCAGCTGCGAAG 540  
3157 CAACCTGTGGGAAGGCGATGCGTGCAGGCGCTTCTTCCCTAATTAAGCCAGCTGCGAAG 3098  
541 GGGATGTGCTGCAAGGCGATTAAGTGGGTAAACGCAAGGTTTTCCGACGTACGACGCTG 600  
3097 GGGATGTGCTGCAAGGCGATTAAGTGGGTAAACGCAAGGTTTTCCGACGTACGACGCTG 3038  
601 TAAACGACGCGCAAGTGAAGCGCGCTCGTCAATTCACGTTTTTGAACCGGTGAGAGAG 660  
3037 TAAACGACGCGCAAGTGAAGCGCGCTCGTCAATTCACGTTTTTGAACCGGTGAGAGTCA 2978  
661 GCAAGCTGCGGTCGCAATGTGTTTTACAGCGTATGAGCAAGATGAAGATGCTGACAC 720  
2977 CCGGCGTGGCGG-----CCGCTAAGAACTAAGTGAATCCCG--CGG 2939  
721 GCTGCAAGACGCGACGTAAGTAAACCTAAGAAAGTAAATCAATGTGAGAGTACGTTAA 780  
2938 GCTGCAAGAACTTCTTAATTTGATGTTATCATGTCACATGTGTATTCACACAGAGTTA 2879  
781 AGATATATCATGCTAATTAATGACGATGGAATCTGTATATACAGCTCACTATAGGCGCAT 840  
2878 CAATCTCAAGAAAGACATGTGTCACGCTTTTGTGGAGTCTTTGCAAAAGGCAAGATT 2819  
841 TGGGTAACGGGCGCGCGCTCGAGGTGACGCTATGCAATGAGCTTGATATCGAATTCCTGC 900  
2818 GTGTCGACAGGTAAATGTGTTCTGTGTAAAGACAGGCGCATTCGCAATGGAATTTT 2759  
901 AGCCGAGGAGATCACTAGTTCTAGAGCGCGCGACCGCGGTGAGAGTCCAGCTTTTGT 960  
2758 GTGATATATGCTGTGATGTTGAACGATTCATCTTCAATGTTGGGGAATTTTGAAGT 2699  
961 TCCCTTATAGTGGGTTAATTAATGATCCCATGCTCAATTTTACGACGACTATCTTTTACG 1020  
2698 TAGCTTATTCATTTCTTTTGTGTTGCTGCGCGTATGTATACATTTGTGTAG--TTATAG 2640

1021 GCTTAATCTACGTGATAGATCATATCGTGGGCTCTTTTTCGCGCTCAGTATCGCC 1080  
2639 TTGTAATCTGAGTGTGTGTCGGAATGTTCCTATCTTTTAAATCAATCTTTTAAAC 2580  
1081 CAAGCTGCGCTATCTGGGCAATCGGGGAGAAAGCCGAGCTTTTCCCGGAGTTG 1140  
2579 TCGATAGCATTAACAAGGATTCACCTTCAAACTTGACCTTGACGACGCGCTTTGTAGTTC 2520  
1141 AAGCGGCAATGAAAGAGTTTCCGAGGATGACTGCTGCTGATGACGTTGAGCGAANAAC 1200  
2519 CCGTATCTTTGAAAGATATAGTGCCTTCGTAATTAACCTTGGGAGATGACCTGTG 2460  
1201 GCAGCTTACCATGATGATTCGGAAGGTGTCGCAATGACGCTTTAATCGGTGAATGT 1260  
2459 AAAAAGTATCCCTTCAATATGATCCGGAATACGGAAGAAAGCATGGAACATTAAGAG 2400  
1261 TCGTTACAGGCAACCGGGAATACAGTTCTGCGGCTTTTCGGAACAAGTTCCGGATG 1320  
2399 AAAGTATGACATGATGTTGGCCATGGAACAGTATGTTTCCAGTATGCAATTAATTA 2340  
1321 TCAACCGGAAGCGATACGAACCGGAACAATAACCGGCAACAGCGGAACTGCGGTGCG 1380  
2339 AGGTAAGCTTTCGTAATGTATGATCACTTCACTTCCTCCTCAGACAGAAATTTGTGC 2280  
1381 GTGTGCAATTAATGACAGCGGTGCGCGCTGGGATTTATAGTCAAGGAGAGGATATC 1440  
2279 CCAATTAATCATCAATCTAATTAATCAAGAAATTTGGGCAACTCCAGTGAATAAGTCTCT 2220  
1441 CTGGCTGGAATGCGGCAAAATGSAATGSAATACCCCGAGTTACCGGCGGGGCGCTT 1500  
2219 C---CTTTACTCATGCGTACCCAGCTTTTGTCTTCTTATGAGAGGTTAATTTGCGCGCTT 2163  
1501 GCGGTAATCATAGTATCATAGCTTTTCTGTGTGAATTTGTATCCGCTCAATTTCAACA 1560  
2162 GCGGTAATCATAGTATCATAGCTTTTCTGTGTGAATTTGTATCCGCTCAATTTCAACA 2103  
1561 CAATATACAGCGCGGAAGCATTAAGTGAAGCTTGGGTCCTTAATGATGATGACTAAT 1620  
2102 CAATATACAGCGCGGAAGCATTAAGTGAAGCTTGGGTCCTTAATGATGATGACTAAT 2043  
1621 CACATTAATGCGTTGCGCTCATAGCTGCGCGCTTCCAGTGGGGAACCTGTGTCGCAAGT 1680  
2042 CACATTAATGCGTTGCGCTCATAGCTGCGCGCTTCCAGTGGGGAACCTGTGTCGCAAGT 1983  
1681 GCATTAATGATCGGCAACGCGCGGAGAGGCGGTTTGGCTATTTGAGCGCTCTTCGCG 1740  
1982 GCATTAATGATCGGCAACGCGCGGAGAGGCGGTTTGGCTATTTGAGCGCTCTTCGCG 1923  
1741 TTCTGCTCATGATGCTGCTGCGCTGCTGTGCTGCTGCGGAGCGGATATCACTCA 1800  
1922 TTCTGCTCATGATGCTGCTGCGCTGCTGTGCTGCTGCGGAGCGGATATCACTCA 1863  
1801 CTCAAGCGGTAATAGGTTATTCACAGAAATCAGGGAATTAACGAGAAAGATCATGTG 1860  
1862 CTCAAGCGGTAATAGGTTATTCACAGAAATCAGGGAATTAACGAGAAAGATCATGTG 1803  
1861 AGCAAAAGGCGCAGCAAAAGCGCAAGACCGTAAAGAGCGCGCTTGTGCTGCTTTTCA 1920  
1802 AGCAAAAGGCGCAGCAAAAGCGCAAGACCGTAAAGAGCGCGCTTGTGCTGCTTTTCA 1743  
1921 TAGGCTCGCGCGCTTACAGACATCACAAAATTCAGCTCAAGTCAAGAGTGGGCAAA 1980  
1742 TAGGCTCGCGCGCTTACAGACATCACAAAATTCAGCTCAAGTCAAGAGTGGGCAAA 1683  
1981 CCGGACAGGATTAATTAAGATACAGAGGCTTTCCCGCTGGAAGCTCCGTGCGCTTCC 2040  
1682 CCGGACAGGATTAATTAAGATACAGAGGCTTTCCCGCTGGAAGCTCCGTGCGCTTCC 1623  
2041 TGTTCGACCTGCGCTTACCGGATACCTGTGCGCTTTCTCCTTTCGGAAGCGTGGC 2100  
1622 TGTTCGACCTGCGCTTACCGGATACCTGTGCGCTTTCTCCTTTCGGAAGCGTGGC 1563

QY 2101 GCTTTCATAGTCAAGCTGTAAGTATCTAGTTCAGTTCGATGAGGTCGTCCTCAAGCT 2160  
 Db 1562 GCTTTCATAGTCAAGCTGTAAGTATCTAGTTCAGTTCGATGAGGTCGTCCTCAAGCT 1503  
 QY 2161 GGGCTGTGTGCAAGAACCCCGTTCAGCCGACCGGCTTATCCGTAACATATCG 2220  
 Db 1502 GGGCTGTGTGCAAGAACCCCGTTCAGCCGACCGGCTTATCCGTAACATATCG 1443  
 QY 2221 TCTTGAAGTCAACCCGTAAGACACATTAATGCGCACTGGCAGCAGCAGTGTAAACAG 2280  
 Db 1442 TCTTGAAGTCAACCCGTAAGACACATTAATGCGCACTGGCAGCAGCAGTGTAAACAG 1383  
 QY 2281 GATTAGCAGACGAGATGTAGGCGGTGCTACAGAGTCTTGAAGTGTGCTTAACTA 2340  
 Db 1382 GATTAGCAGACGAGATGTAGGCGGTGCTACAGAGTCTTGAAGTGTGCTTAACTA 1323  
 QY 2341 CGGCTACACTAGAAAGACATATTGGTATCTGCGCTGTGTAAGCAGATCTTGAATCTT 2400  
 Db 1322 CGGCTACACTAGAAAGACATATTGGTATCTGCGCTGTGTAAGCAGATCTTGAATCTT 1263  
 QY 2401 AAAAAGATTGTAAGTCTTGAATCTGCGCAACAAACACCGTGTGCGGTGTTT 2460  
 Db 1262 AAAAAGATTGTAAGTCTTGAATCTGCGCAACAAACACCGTGTGCGGTGTTT 1203  
 QY 2461 TGTTCGAAAGCAGACATTAACGCGCAAGAAAGATCTCAAGAGATCTTGAATCTT 2520  
 Db 1202 TGTTCGAAAGCAGACATTAACGCGCAAGAAAGATCTCAAGAGATCTTGAATCTT 1143  
 QY 2521 TTTCAAGGGGTGTAGGCTGCTAGTGAAGAAACCTCAAGTAAAGGATCTTGAATCTT 2580  
 Db 1142 TTTCAAGGGGTGTAGGCTGCTAGTGAAGAAACCTCAAGTAAAGGATCTTGAATCTT 1083  
 QY 2581 ATTATCAAAAGATCTTCAAGTCTTGAATCTTGAATCTTGAATCTTGAATCTT 2640  
 Db 1082 ATTATCAAAAGATCTTCAAGTCTTGAATCTTGAATCTTGAATCTTGAATCTT 1023  
 QY 2641 CTAAGTATATATAGTAACTGCTGCTAGCAGTAACTCAATCTTGAATCTTGAATCTT 2700  
 Db 1022 CTAAGTATATATAGTAACTGCTGCTAGCAGTAACTCAATCTTGAATCTTGAATCTT 963  
 QY 2701 TATCTCAGCATCTGCTATTTGCTATCTATCTATCTATCTATCTATCTATCTATCTATCT 2760  
 Db 962 TATCTCAGCATCTGCTATTTGCTATCTATCTATCTATCTATCTATCTATCTATCTATCT 903  
 QY 2761 AACTACGATACGGAAGGCTTACACATCTGCGCCAGAGCTGCAATGATATCCGGAACCC 2820  
 Db 902 AACTACGATACGGAAGGCTTACACATCTGCGCCAGAGCTGCAATGATATCCGGAACCC 843  
 QY 2821 ACCTCAGCGGCTCCAGATTTATACGAAATTAACACAGCAGCGGAAGGCGAGCGAG 2880  
 Db 842 ACCTCAGCGGCTCCAGATTTATACGAAATTAACACAGCAGCGGAAGGCGAGCGAG 783  
 QY 2881 AAGTGTCTGCAACTTATCTGCTCCTATCTATCTATCTATCTATCTATCTATCTATCT 2940  
 Db 782 AAGTGTCTGCAACTTATCTGCTCCTATCTATCTATCTATCTATCTATCTATCTATCT 723  
 QY 2941 AGTAAAGTATGCGCACTTAATAGTTTGGCAAGTCTTGTGCAATCTGCAAGCATCTG 3000  
 Db 722 AGTAAAGTATGCGCACTTAATAGTTTGGCAAGTCTTGTGCAATCTGCAAGCATCTG 663  
 QY 3001 GGTGTACGCTCGTGTGTTGTTGCTTCAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3060  
 Db 662 GGTGTACGCTCGTGTGTTGTTGCTTCAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 603  
 QY 3061 AGTTACATGATCCCGCATGTTGCAAAAAAGCGTTAGTCTTCTGCTGCTGCTGCTGCT 3120  
 Db 602 AGTTACATGATCCCGCATGTTGCAAAAAAGCGTTAGTCTTCTGCTGCTGCTGCTGCT 543  
 QY 3121 TGTCAAAATTAAGTTGGCGCAGTGTATCACTAGTTATGCGCAGCATCTGCAATATTC 3180  
 Db 542 TGTCAAAATTAAGTTGGCGCAGTGTATCACTAGTTATGCGCAGCATCTGCAATATTC 483  
 QY 3181 TCTTACTGTCATGCAATCCGTAAATGCTTTTCTGTGACTGTGATCTCAACCAAGTC 3240

Db 482 TCTTACTGTCATGCAATCCGTAAATGCTTTTCTGTGACTGTGATCTCAACCAAGTC 423  
 QY 3241 ATTCTAGAAATAGTGTATGTCGCGCAGCAGAGTGTCTTGGCGGCTCAATACGGAATTA 3300  
 Db 422 ATTCTAGAAATAGTGTATGTCGCGCAGCAGAGTGTCTTGGCGGCTCAATACGGAATTA 363  
 QY 3301 TACCGCGCAGATGACAGAACTTTAAAGTGTCTCATCTATGGAACAAAGTTCTTGGGCG 3360  
 Db 362 TACCGCGCAGATGACAGAACTTTAAAGTGTCTCATCTATGGAACAAAGTTCTTGGGCG 303  
 QY 3361 AAACTCTCAAGATCTTCAACCGCTGTGAGATCTCAAGTGTGAATCCCACTCGGTGAC 3420  
 Db 302 AAACTCTCAAGATCTTCAACCGCTGTGAGATCTCAAGTGTGAATCCCACTCGGTGAC 243  
 QY 3421 CAACGTATCTTCAAGATCTTCTTCACTTCAACGCTTCTGCGGTGAGCAAAACAGGAG 3480  
 Db 242 CAACGTATCTTCAAGATCTTCTTCACTTCAACGCTTCTGCGGTGAGCAAAACAGGAG 183  
 QY 3481 GCAAAATGCGCAGAAAGGAAATAGGCGCAGCAGAAATGTTGAATCTCAATCTCTT 3540  
 Db 182 GCAAAATGCGCAGAAAGGAAATAGGCGCAGCAGAAATGTTGAATCTCAATCTCTT 123  
 QY 3541 CCTTTTCAATTTATTTGAAGATTTATCAAGGTTATTTGCTCATAGAGGATCAATTT 3600  
 Db 122 CCTTTTCAATTTATTTGAAGATTTATCAAGGTTATTTGCTCATAGAGGATCAATTT 63  
 QY 3601 TGAATGTTATTTGAATTTAAATTAACAAATAGGCGTTCGCGCAGTTCCTCCGAAAGTCC 3660  
 Db 62 TGAATGTTATTTGAATTTAAATTAACAAATAGGCGTTCGCGCAGTTCCTCCGAAAGTCC 3  
 QY 3661 AC 3662  
 Db 2 AC 1

RESULT 13  
 US-10-684-134-17/c  
 ; Sequence 17, Application US/10684134  
 ; Publication No. US2004014243A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Large Scale Biology Corporation  
 ; TITLE OF INVENTION: POPULATION OF POLYNUCLEOTIDE SEQUENCE VARIANTS  
 ; FILE REFERENCE: LSBC-PLG5392-CIP  
 ; CURRENT APPLICATION NUMBER: US/10/684,134  
 ; PRIOR FILING DATE: 2003-10-10  
 ; PRIOR APPLICATION NUMBER: 60/402,342  
 ; PRIOR FILING DATE: 2002-08-08  
 ; PRIOR APPLICATION NUMBER: 10/066,390  
 ; PRIOR FILING DATE: 2002-02-01  
 ; PRIOR APPLICATION NUMBER: 60/268,785  
 ; PRIOR FILING DATE: 2001-02-14  
 ; PRIOR APPLICATION NUMBER: 60/266,386  
 ; PRIOR FILING DATE: 2002-02-02  
 ; NUMBER OF SEQ ID NOS: 35  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 17  
 ; LENGTH: 3637  
 ; TYPE: DNA  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Encodes cycle 3 GFP  
 US-10-684-134-17

Query Match 72.8%; Score 2664.8; DB 17; Length 3637;  
 Best Local Similarity 83.8%; Pred. No. 0;  
 Matches 3070; Conservative 0; Mismatches 567; Indels 25; Gaps 4;  
 QY 1 CTTAAATTTGAAGCGTTAATTTTGTAAATTCGCTTAAATTTTGTAAATCAAGTC 60  
 Db 3637 CTTAAATTTGAAGCGTTAATTTTGTAAATTCGCTTAAATTTTGTAAATCAAGTC 60  
 QY 61 ATTTTAAACCAATAGCCGCAAAATCGCAAAATCCCTTATTAATCAAAAGATGACCGA 120

Db 3577 ATTTTAAACCAATAGCCGAAATCGCAAAATCCCTTTAATCAAAAGATTAACCGA 3518  
| | | | |  
Qy 121 GATAGGGTTGAGTGTGTTCCAGTTTGGAAACAAGATCCATTTAAAGAGCTGGAATC 180  
| | | | |  
Db 3517 GATAGGGTTGAGTGTGTTCCAGTTTGGAAACAAGATCCATTTAAAGAGCTGGAATC 3458  
| | | | |  
Qy 181 CAAGCTCAAGGGCCGAAAAACCGTCTATCAGGGCGATGGCCCACTAAGTGAACCTACAC 240  
| | | | |  
Db 3457 CAAGCTCAAGGGCCGAAAAACCGTCTATCAGGGCGATGGCCCACTAAGTGAACCTACAC 3398  
| | | | |  
Qy 241 CTAATCAAGTTTGTGGGTCGAGGTGCGTAAAGCACTAAATCGGAACCTTAAGGGAG 300  
| | | | |  
Db 3397 CTAATCAAGTTTGTGGGTCGAGGTGCGTAAAGCACTAAATCGGAACCTTAAGGGAG 3338  
| | | | |  
Qy 301 CCCCCGATTTAGAGCTTGAACGGGGAAACCGGCGAAACGTGGCGAAGAAAGAGGAGAA 360  
| | | | |  
Db 3337 CCCCCGATTTAGAGCTTGAACGGGGAAACCGGCGAAACGTGGCGAAGAAAGAGGAGAA 3278  
| | | | |  
Qy 361 AGCCAAAGAGACGGGCGCTTAGGGCGCTGAGAGTAAAGCTGACAGCTGCGGTAAACAC 420  
| | | | |  
Db 3277 AGCCAAAGAGACGGGCGCTTAGGGCGCTGAGAGTAAAGCTGACAGCTGCGGTAAACAC 3218  
| | | | |  
Qy 421 CACACCCGCGCTTAAATGCGCCGCTACAGGGCGCGTCCCATTCGCAATTCAGGCTGCG 480  
| | | | |  
Db 3217 CACACCCGCGCTTAAATGCGCCGCTACAGGGCGCGTCCCATTCGCAATTCAGGCTGCG 3158  
| | | | |  
Qy 481 CAAGCTTTGGAAAGGCGATGCTGCGGCGCTCTTGCTAATTAACCGCAGCTGGCGAAGG 540  
| | | | |  
Db 3157 CAAGCTTTGGAAAGGCGATGCTGCGGCGCTCTTGCTAATTAACCGCAGCTGGCGAAGG 3098  
| | | | |  
Qy 541 GGGATGCTGCTCAAGGCGCATTAAGTTGGGTAAACGCAAGGTTTTCCAGTCAACGATG 600  
| | | | |  
Db 3097 GGGATGCTGCTCAAGGCGCATTAAGTTGGGTAAACGCAAGGTTTTCCAGTCAACGATG 3038  
| | | | |  
Qy 601 TAAAAAGAGGCGCAGTGAAGCGCGCTCGTTCAATTAACGTTTTTGAACCGGTGAGAGAG 660  
| | | | |  
Db 3037 TAAAAAGAGGCGCAGTGAAGCGCGCTCGTTCAATTAACGTTTTTGAACCGGTGAGAGAG 2978  
| | | | |  
Qy 661 GCAAGCTGCGGCTGCAAAATGTGTTTTACAGCGTGAATGAGCAAGATGCTGACAC 720  
| | | | |  
Db 2977 CCGCGGTGGCGG-----CGGCTTAAGAACTGATGATCCCG--CGG 2939  
| | | | |  
Qy 721 GCTCGAAGAACGCAAGCTAAGTAAACCTAAGAAATTAATATATTGACAGTAACTTTAA 780  
| | | | |  
Db 2938 GCTCGAAGAAATTTATATTGTAATGATTATCATATCCATGTAATCCACAGCAGATTA 2879  
| | | | |  
Qy 781 AGATTAATCATGCTAATAATTGACGATGGGATCTGTAAATAGACCTCACTAATAGGCGAAT 840  
| | | | |  
Db 2878 CAAACTCAAGAAAGACATGTGTGTCACGCTTTTGTGGGATCTTTGCAAAAGGCGAAT 2819  
| | | | |  
Qy 841 TGGGTACGGGCGCCCGCTGAGGTGCAAGGTATGATAGCTGATATCGAATTCCTG 900  
| | | | |  
Db 2818 GTGTGACAGGTAATGTGTCTGGTAAAGAACAGGGCCATTCGCAATTTGAGATTTT 2759  
| | | | |  
Qy 901 AGCCCGGGGATTCACATGTCTTAGAGCGCGCGCACCGCGTGAAGTCCAGCTTTGT 960  
| | | | |  
Db 2758 GTTGAATATGCTGTGATGTGAACGATTCATCTTCAATGTGTGGGAAATTTTGAAGT 2699  
| | | | |  
Qy 961 TCCCTTAGTAGAGGTTAATTAGATCCCATGCTCAATTTTACGACATATCTTTAG 1020  
| | | | |  
Db 2698 TAGCTTTGATTCATCTTTGTTTGTGTCGCGTGAATATATGTTGATAG--TTATAG 2640  
| | | | |  
Qy 1021 GATTATCTAGCTGATCAGATCAATATCGTGGGTCTTTTTCGCGCTCAGTATCGCC 1080  
| | | | |  
Db 2639 TTGTACTGAGATTGTGTCCGAGATGTTTCCATCTTCTTAATAATCAATCTTTTAA 2580  
| | | | |  
Qy 1081 CAAGCTGCGCTATCTGTGGCATTCGGGAGAGAAAGCCGCTGCTTTCCGCGAGGTG 1140  
| | | | |  
Db 2579 TCGATAGCATTTAAAGGGATTCACCTTCAAACTTGATCTTCAGACGCGCTTTGTAGTTC 2520  
| | | | |  
Qy 1141 AAGCGGATGGAAGAGTTTGCAGAGATGACTGCTGTGATGACGTTGAGGAGAAAC 1200  
| | | | |

Db 2519 CCGTCACTTTGAAGAATATAGTGCCTTCTGTACATTAACCTTGGGAGTGGCACTCTTG 2460  
| | | | |  
Qy 1201 GCAGCTTTACCATGATGATTCGGGAAGGTGTGCCATATGACAGCCCTTTAAGGTGAATCT 1260  
| | | | |  
Db 2459 AAAAAGTCAATCCGTTTCAATATGATCCGGAATACGGGAAAGCATTTGAACCATTAAG 2400  
| | | | |  
Qy 1261 TCGTTCAAGCCACCTGGGATACAGATTGCTGTGCGGCTTTTCCGACACAGATTCGGAATG 1320  
| | | | |  
Db 2399 AAAGTATGACAAAGTTGGTGGCCATGGAACAGTATGTTTTCCAGTATGCAATAATTTA 2340  
| | | | |  
Qy 1321 TCAGCCCGAAGCGATCAGCAACCGAACAATATCCGCGCAAGCGCGAATCGCGTGGCG 1380  
| | | | |  
Db 2339 AGGTTAAGCTTTCCGTAATGATGATCACTTCAACCTCTCTCACTGACAGAAATTTGTGC 2280  
| | | | |  
Qy 1381 GTGTGACATTAATGACAGCGGTGCGGCTGGGATATTACGTACAGCGAGACGGGTATC 1440  
| | | | |  
Db 2279 CCATTAACATACCATCTAATTTCAACAAGATTTGGACAACTCCAGTGAATAAGTTCTTCT 2220  
| | | | |  
Qy 1441 CTGAGTGAATGCCCAATAATGACATGGAATACCCCGTAGTACCCGCGGCGGCTT 1500  
| | | | |  
Db 2219 C--CTTTACTCATCGGTACCCAGCTTTTGTCTTTAGTGAAGGTTAATTTGGCGCTT 2163  
| | | | |  
Qy 1501 GCGCTAATCATAGTCAATAGCTGTTTCTGTGTGAATTTGTAATCGCTCAATTTCCACA 1560  
| | | | |  
Db 2162 GCGCTAATCATAGTCAATAGCTGTTTCTGTGTGAATTTGTAATCGCTCAATTTCCACA 2103  
| | | | |  
Qy 1561 CAACATACAGACCGGAAGCATTAAGTAAAGCTTGGGGTGCCTAATGATGATTAATCT 1620  
| | | | |  
Db 2102 CAACATACAGACCGGAAGCATTAAGTAAAGCTTGGGGTGCCTAATGATGATTAATCT 2043  
| | | | |  
Qy 1621 CACATTAATTCGTTGGGCTCACTGCGCCCTTTTCACTCGGAAACCTGTGTGTCCAGCT 1680  
| | | | |  
Db 2042 CACATTAATTCGTTGGGCTCACTGCGCCCTTTTCACTCGGAAACCTGTGTGTCCAGCT 1983  
| | | | |  
Qy 1681 GCATTAATGAAATCGGCCAAACGCGGGGAGAGGCGTTTGTGTTTGGGCGCTCTTCGCG 1740  
| | | | |  
Db 1982 GCATTAATGAAATCGGCCAAACGCGGGGAGAGGCGTTTGTGTTTGGGCGCTCTTCGCG 1923  
| | | | |  
Qy 1741 TTCCTGCTCACTGACTCGCTGCGCTGCTGCTTCTGCTGCGGCGAGCGGTATGACTCA 1800  
| | | | |  
Db 1922 TTCCTGCTCACTGACTCGCTGCGCTGCTGCTGCTGCTGCGGCGAGCGGTATGACTCA 1863  
| | | | |  
Qy 1801 CTCAAAAGCGGTTAATACGTTATTCACAGATTCAGGGGATTAACGCAAGAAAGAACTGTG 1860  
| | | | |  
Db 1862 CTCAAAAGCGGTTAATACGTTATTCACAGATTCAGGGGATTAACGCAAGAAAGAACTGTG 1803  
| | | | |  
Qy 1861 AGCAAAAGGCGCAGCAAAAGGCGAGAAACCGTAAAGAGCGCGCTGTGGGCTTTTCCA 1920  
| | | | |  
Db 1802 AGCAAAAGGCGCAGCAAAAGGCGAGAAACCGTAAAGAGCGCGCTGTGGGCTTTTCCA 1743  
| | | | |  
Qy 1921 TAGGCTCGGCGCCCTGACGAGCATCACAAAATGACGCTCAAGTCAAGAGGTGCGAA 1980  
| | | | |  
Db 1742 TAGGCTCGGCGCCCTGACGAGCATCACAAAATGACGCTCAAGTCAAGAGGTGCGAA 1683  
| | | | |  
Qy 1981 CCGCAGAGCATTAAGATTAACAGAGGCTTTCCCGCTGGAAGCTCCTCGTGCCTTCC 2040  
| | | | |  
Db 1682 CCGCAGAGCATTAAGATTAACAGAGGCTTTCCCGCTGGAAGCTCCTCGTGCCTTCC 1623  
| | | | |  
Qy 2041 TGTTCGAGCCCTGCGCTTAACCGGATACCTGTGCGCTTTCTCCCTGCGGAAGCGTGGC 2100  
| | | | |  
Db 1622 TGTTCGAGCCCTGCGCTTAACCGGATACCTGTGCGCTTTCTCCCTGCGGAAGCGTGGC 1563  
| | | | |  
Qy 2101 GCTTTCATATAGCTCAAGCTGTAGTATCTCAATTCGATTCGATTCGCTTCAAGCT 2160  
| | | | |  
Db 1562 GCTTTCATATAGCTCAAGCTGTAGTATCTCAATTCGATTCGATTCGCTTCAAGCT 1503  
| | | | |  
Qy 2161 GGGCTGTGTGACAGAAACCCCGCTTACGCGCCGACCGTGGCTTATTCGGTAACTATG 2220  
| | | | |  
Db 1502 GGGCTGTGTGACAGAAACCCCGCTTACGCGCCGACCGTGGCTTATTCGGTAACTATG 1443  
| | | | |  
Qy 2221 TCTTAGTCAACCCGCTAAGACACGACTATGCGCCCTGAGACGACGCACTGTTAAG 2280  
| | | | |  
Db 1442 TCTTAGTCAACCCGCTAAGACACGACTATGCGCCCTGAGACGACGCACTGTTAAG 1383  
| | | | |



QY 2281 GATTAGCAGAGCGAGGATGTAGACGGTCTGACAGAGTCTTGAAGTGGTCCATACTA 2340  
DB 1382 GATTAGCAGAGCGAGGATGTAGACGGTCTGACAGAGTCTTGAAGTGGTCCATACTA 1323  
QY 2341 CGGCTACACTAGAGAGCAGATATTTGTATCTGCGCTGCTGCGAAGCCAGTTACCTTCGG 2400  
DB 1322 CGGCTACACTAGAGAGCAGATATTTGTATCTGCGCTGCTGCGAAGCCAGTTACCTTCGG 1263  
QY 2401 AAAAAAGATGTAGAGTCTTGAATCCGGCAACAAACACCGCTGTAGCGGTGTTT 2460  
DB 1262 AAAAAAGATGTAGAGTCTTGAATCCGGCAACAAACACCGCTGTAGCGGTGTTT 1203  
QY 2461 TGTTCGACAGCAGAGATTACGCGAGAAAAAGAGATCTCAAGATCTTTGATCTT 2520  
DB 1202 TGTTCGACAGCAGAGATTACGCGAGAAAAAGAGATCTCAAGATCTTTGATCTT 1143  
QY 2521 TTTCAAGGGGTCTGACGCTCAGTGGAGAAAACTCAAGTTAAGGGATTTTGTGATAG 2580  
DB 1142 TTTCAAGGGGTCTGACGCTCAGTGGAGAAAACTCAAGTTAAGGGATTTTGTGATAG 1083  
QY 2581 ATTATCAAAAAAGATCTTCACTAGATCTTTTAAATTAATTAAGATTTTAATCAAT 2640  
DB 1082 ATTATCAAAAAAGATCTTCACTAGATCTTTTAAATTAATTAAGATTTTAATCAAT 1023  
QY 2641 CTTAAGTATATATAGTAACTTGTCTGACAGTTACCAATGCTTAATCAGTAGGCAAC 2700  
DB 1022 CTTAAGTATATATAGTAACTTGTCTGACAGTTACCAATGCTTAATCAGTAGGCAAC 963  
QY 2701 TATCTCAGCGATCTGTCTATTTGTTTCATCATCAGTATGCTGATCTCCGTGTGTAGAT 2760  
DB 962 TATCTCAGCGATCTGTCTATTTGTTTCATCATCAGTATGCTGATCTCCGTGTGTAGAT 903  
QY 2761 AACTAGCATGCGGAGGGGCTTACACTCTGCGCCAGTGTGCAATGATACCGGAGACC 2820  
DB 902 AACTAGCATGCGGAGGGGCTTACACTCTGCGCCAGTGTGCAATGATACCGGAGACC 843  
QY 2821 ACCGTCAACGGGCTCCAGATTATACAGCAATTAACAGCAGCCGGAAGGGCCGACG 2880  
DB 842 ACCGTCAACGGGCTCCAGATTATACAGCAATTAACAGCAGCCGGAAGGGCCGACG 783  
QY 2881 AAGTGTCTCCGCAACTTTATCCGCTCCATCCAGTCTTAATTTGTTGCGGGAAGCTAG 2940  
DB 782 AAGTGTCTCCGCAACTTTATCCGCTCCATCCAGTCTTAATTTGTTGCGGGAAGCTAG 723  
QY 2941 AGTAAGTATGCGCAGTTAATAGTTTSGCAAGGTTGTTGCCATTCCTACAGCATCTG 3000  
DB 722 AGTAAGTATGCGCAGTTAATAGTTTSGCAAGGTTGTTGCCATTCCTACAGCATCTG 663  
QY 3001 GGTGTCAAGCTGTGCTTGTATGCTTCATTGAGCTCCGTTCCCAAGCATCAAGGGG 3060  
DB 662 GGTGTCAAGCTGTGCTTGTATGCTTCATTGAGCTCCGTTCCCAAGCATCAAGGGG 603  
QY 3061 AGTTATGATGATCCCCCATGTTGTGCAAAAAAGGGTTAGTCTTGTGCTTCCTCCATCTG 3120  
DB 602 AGTTATGATGATCCCCCATGTTGTGCAAAAAAGGGTTAGTCTTGTGCTTCCTCCATCTG 543  
QY 3121 TGTGAGAGTATGTTGCGCAGTGTATATCACTCATGTTATGCGACAGCTCATATATTC 3180  
DB 542 TGTGAGAGTATGTTGCGCAGTGTATATCACTCATGTTATGCGACAGCTCATATATTC 483  
QY 3181 TCTTACTGTATGATCCGTAAGATGCTTTCTGTGAGTGTGTGATCTCAACCAAGTC 3240  
DB 482 TCTTACTGTATGATCCGTAAGATGCTTTCTGTGAGTGTGTGATCTCAACCAAGTC 423  
QY 3241 ATTCTGAGAAATATGATGCGGACCGAGTTGCTTGTCCCGGCGTCAATAGCGGATTA 3300  
DB 422 ATTCTGAGAAATATGATGCGGACCGAGTTGCTTGTCCCGGCGTCAATAGCGGATTA 363  
QY 3301 TACCGGCGCACTAGAGCACTTTAAAGTGTCTCATCTGGAAGAAAGTTCTTGGGGCG 3360  
DB 362 TACCGGCGCACTAGAGCACTTTAAAGTGTCTCATCTGGAAGAAAGTTCTTGGGGCG 303

QY 3361 AAAACTCAAGAGATCTTACCGCTGTGAATCCAGTTGCATGTAAACCACTGTCACCC 3420  
DB 302 AAAACTCAAGAGATCTTACCGCTGTGAATCCAGTTGCATGTAAACCACTGTCACCC 243  
QY 3421 CAACGTATCTTACAGATCTTTTACTTTCACACAGGTTCTGAGGTGAGCAAAAAAGAG 3480  
DB 242 CAACGTATCTTACAGATCTTTTACTTTCACACAGGTTCTGAGGTGAGCAAAAAAGAG 183  
QY 3481 GCAAAATGCGCCAAAGGAAATAGGGGAGACCGAAATGTTGAATCTCATCTCTT 3540  
DB 182 GCAAAATGCGCCAAAGGAAATAGGGGAGACCGAAATGTTGAATCTCATCTCTT 123  
QY 3541 CCTTTTCATTAATTAATGAAGATTTATCAGGGTTATGTCTCATGAGCGGATACATATT 3600  
DB 122 CCTTTTCATTAATTAATGAAGATTTATCAGGGTTATGTCTCATGAGCGGATACATATT 63  
QY 3601 TGAATGATTTAGAAAAATTAACAAATAGGGGTTCCGCGACATTTCCCGAAAAAGTCC 3660  
DB 62 TGAATGATTTAGAAAAATTAACAAATAGGGGTTCCGCGACATTTCCCGAAAAAGTCC 3  
QY 3661 AC 3662  
DB 2 AC 1

RESULT 14  
US-10-637-758-17/c  
Sequence 17, Application US/10637758  
Publication No. US20040180352A1  
GENERAL INFORMATION:  
APPLICANT: Large Scale Biology Corporation  
TITLE OF INVENTION: POPULATION OF POLYNUCLEOTIDE SEQUENCE VARIANTS  
FILE REFERENCE: LSBC-PIG5392-CIP  
CURRENT APPLICATION NUMBER: US/10/637,758  
PRIOR FILING DATE: 2003-08-08  
PRIOR APPLICATION NUMBER: US/10/280,913  
PRIOR FILING DATE: 2002-10-25  
PRIOR APPLICATION NUMBER: 60/402,342  
PRIOR FILING DATE: 2002-08-08  
PRIOR APPLICATION NUMBER: 10/066,390  
PRIOR FILING DATE: 2002-02-01  
PRIOR APPLICATION NUMBER: 60/268,785  
PRIOR FILING DATE: 2001-02-14  
PRIOR APPLICATION NUMBER: 60/266,386  
PRIOR FILING DATE: 2002-02-02  
NUMBER OF SEQ ID NOS: 35  
SOFTWARE: Patentin version 3.1  
SEQ ID NO 17  
LENGTH: 3637  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Encodes cycle 3 GFP  
US-10-637-758-17

Query Match 72.8%; Score 2664.8; DB 17; Length 3637;  
Best Local Similarity 83.8%; Pred. No. 0;  
Matches 3070; Conservative 0; Mismatches 567; Indels 25; Gaps 4;  
QY 1 CTTAATTTGTAAGCGTTAATATTTTGTAAATTCGCGTAAATTTTGTAAATTCGCTC 60  
DB 3637 CTTAATTTGTAAGCGTTAATATTTTGTAAATTCGCGTAAATTTTGTAAATTCGCTC 3578  
QY 61 ATTTTAAACCAATAGGCGGAAATCCGCAAAATCCCTTAATTAATCAAAAGATAGACCGA 120  
DB 3577 ATTTTAAACCAATAGGCGGAAATCCGCAAAATCCCTTAATTAATCAAAAGATAGACCGA 3518  
QY 121 GATAGGGTGAAGTGTGTTGTTGCAAGTTGGAACAAGAGTCCACTATTAAAGAGTGAATC 180  
DB 3517 GATAGGGTGAAGTGTGTTGTTGCAAGTTGGAACAAGAGTCCACTATTAAAGAGTGAATC 3458  
QY 181 CAAAGTCAAGAGGCGAAAAACGGTCTATACAGGCGGATGCGCACTAGCTGAACATCACC 240



Db 3457 CAACGTCAAGGGGCGAAAAACGCTCTATCAGGGCGATGGCCCACTACGTAACATCAACC 3398  
Oy 241 CTAATCAAGTTTTTGGGGTGGAGTCCGTTAAACACTAAATCGAAACCTTAAAGGAG 300  
Db 3397 CTAATCAAGTTTTTGGGGTGGAGTCCGTTAAACACTAAATCGAAACCTTAAAGGAG 3338  
Oy 301 CCCCCGATTTAGAGCTTGAACGGGGAAAGCCGGCGAAACGTGGCGAAGAAAGGAGAGAA 360  
Db 3337 CCCCCGATTTAGAGCTTGAACGGGGAAAGCCGGCGAAACGTGGCGAAGAAAGGAGAGAA 3278  
Oy 361 AGCGAAAGAGACGGGCGCTAGGCGCTGCAAGTGTAGCGGTCAAGCTGCGGTAAACAC 420  
Db 3277 AGCGAAAGAGACGGGCGCTAGGCGCTGCAAGTGTAGCGGTCAAGCTGCGGTAAACAC 3218  
Oy 421 CACACCCCGCGCTTAAATGGCCGCTACAGGGCCGCTCCCATTTGCCATTCAAGCTGCG 480  
Db 3217 CACACCCCGCGCTTAAATGGCCGCTACAGGGCCGCTCCCATTTGCCATTCAAGCTGCG 3158  
Oy 481 CAACCTGTTGGAAAGGCGATCGGTGCGGGCCCTTCCGTTAATGACGACGCTGGCGAAG 540  
Db 3157 CAACCTGTTGGAAAGGCGATCGGTGCGGGCCCTTCCGTTAATGACGACGCTGGCGAAG 3098  
Oy 541 GGGATGTGCTCAAGGCGATTAAGTTGGTAAACGCGAGGTTTTTCCAGTCAAGACGCTTG 600  
Db 3097 GGGATGTGCTCAAGGCGATTAAGTTGGTAAACGCGAGGTTTTTCCAGTCAAGACGCTTG 3038  
Oy 601 TAAAAAGACGGGCGATGAGCCGCTCGTTCACTTCAAGTTTTGAACCCGTGGAGAGACGG 660  
Db 3037 TAAAAAGACGGGCGATGAGCCGCTCGTTCACTTCAAGTTTTGAACCCGTGGAGACCTCA 2978  
Oy 661 GCAGACTGCGGATGCAAAATGTGTTTACAGCGGTATGAGCAGATGAAGATGCTGACAC 720  
Db 2977 CCGGCGTGGCGG-----CCGCTTAAAGCTAAGTGAATCCG--CGG 2939  
Oy 721 GCTCGAAGAACGCGAGCTAGTATACCTTAAAGAAATATCATATTGTGACGTAAGTTAA 780  
Db 2938 GCTCGAAGAAATTTCTTATTTGTATAGTTCATCCATGCTCAATGCTGAATCCAGAGACGTTA 2879  
Oy 781 AGATATATATGCTTAAATTTGAACGATGGATCTGTATTAAGACTCACTAATAGGCGCAAT 840  
Db 2878 CAACCTCAAGAGACCAATGTGTGTACACCTTTTCGTTGGGATCTTTCCAAAGGCGAGATT 2819  
Oy 841 TGGGATACGGGCGCCCGCTCGAGGTGACGAGTATGATATGATATGATATTCCTGCG 900  
Db 2818 GTTGCGACAGGTAAATGTTGTCTGTGTAAGAAAGACAGGCGCATCGCAATTTGAAGATT 2759  
Oy 901 AGCCCGGGGATCCACTAGTTCTAGAGCGGCGCCACGCGGTGAGCTCAGACTTTTGT 960  
Db 2758 GTTGAATATGGTCTGTAGTGAACGATCCATCTTCAATGTGTGGCGAATTTGAAGT 2699  
Oy 961 TCCCTTTAGTGAAGGTTAATTAGATCCATGCGTCAATTTTACGCAACTATCTTTTGA 1020  
Db 2698 TAGCTTGAATTCATCTTTGTTGTGTGCGGTGATGATATCAATGTGTAG--TTATAG 2640  
Oy 1021 GGTAAATCTAGCTGATCAGATATATGTCGCGCTTTTCCGCTCAGTCAATCCG 1080  
Db 2639 TTGATCTGAGTTGTGTGCGAAGATGTTTCCATCTTCTTAAATCAATACCTTTTAAC 2580  
Oy 1081 CAAGCTGCGCTATCTGGGCAATCGGGAGAGAAAGCCGTCCTTTTCCGCGAGGTTG 1140  
Db 2579 TCGATATGATTAACAGGGTATCACTTCAAACTTGACTTCAGACAGCGGTCTTTGATGTT 2520  
Oy 1141 AAGCGGATGGAAGAGTTTCCCGAGATGACTGCTGCTGATGAGCTTGAAGGAAAC 1200  
Db 2519 CCGTCATCTTTGAAGATATGTGTGCTCTGTATCAATCACTTGGGCGATGCACTCTTG 2460  
Oy 1201 GCACGTTTACCATGATGATTCGGGAAGGTGTGCGCATGACGCTTTTAAAGGTAACTGT 1260  
Db 2459 AAAAAGTCATGCCCTTTTATATGATCCGATTAACGGGAAGACATTAAGACCAATAGAG 2400  
Oy 1261 TCGTTACAGGCACTGGGATACCAATGCTGCGGCTTTTCCGACCAAGTTCCGAGTGG 1320  
Db 2399 AAAGTAGACAAAGTGTGGCCATGGAACAGGTAGTTTTTCAAGTAGTGCATAAATAATTTA 2340

Oy 1321 TCAGCCGAAAGCGATACGCAACCCGAACATATCCGGGAGACGCCGGAATCGCGTGGC 1380  
Db 2339 AGGGTAAGCTTTCGATATGTGATCACTTACCCCTCTCACTGACGAAATTTGTGCG 2280  
Oy 1381 GTGTGAGATTAATGACAGCGGTGCGGCGGTGATATTAAGTCAAGGAGACGGGTATC 1440  
Db 2279 CCATTAACATCAATCTAATTTCAACAAAGATTTGGACACTCCAGTGAATAATTTCTT 2220  
Oy 1441 CTGGCTGAGTCCGCAAGAAATGACAATGATACCCCGTAGTTACCCGCGGCGCGCTT 1500  
Db 2219 C---CTTTACTCATCGGTAACCAAGCTTTTGTTCCTTTAGTAGAGGTTAATTTGGCGCTT 2163  
Oy 1501 GCGGTATCATGCTGCTATGCTGTTTCTGTGTGAATTTGTTATCCGCTCACAATTTCCA 1560  
Db 2162 GCGGTATCATGCTGCTATGCTGTTTCTGTGTGAATTTGTTATCCGCTCACAATTTCCA 2103  
Oy 1561 CAACATACGACCGGAAGCATPAAAGTGTAAAGCTGGGGTGCCTAATGAGTACCTACT 1620  
Db 2102 CAACATACGACCGGAAGCATPAAAGTGTAAAGCTGGGGTGCCTAATGAGTACCTACT 2043  
Oy 1621 CACATTAATTCGTTGCGCTCACTGCGGCTTTTCCAGTGGGAAACCTGTGTCAGACT 1680  
Db 2042 CACATTAATTCGTTGCGCTCACTGCGGCTTTTCCAGTGGGAAACCTGTGTCAGACT 1983  
Oy 1681 GCATTAATGAATCGGCCCAACGCGCGGGAGAGCGGCTTTGCTATTTGGCGCTCTTCCGC 1740  
Db 1982 GCATTAATGAATCGGCCCAACGCGCGGGAGAGCGGCTTTGCTATTTGGCGCTCTTCCGC 1923  
Oy 1741 TTCTCTGCTCACTGACTGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800  
Db 1922 TTCTCTGCTCACTGACTGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1863  
Oy 1801 CTCAAGCGGTAATACGTTTATCACAAGATCAGGGGATTAACGAGAAAGAACATGTG 1860  
Db 1862 CTCAAGCGGTAATACGTTTATCACAAGATCAGGGGATTAACGAGAAAGAACATGTG 1803  
Oy 1861 AGCAAAAGCGGCAAGAAAGGCGAGAAACCGTAAAGAGCGCGCTTTGCTGCGGTTTTTCCA 1920  
Db 1802 AGCAAAAGCGGCAAGAAAGGCGAGAAACCGTAAAGAGCGCGCTTTGCTGCGGTTTTTCCA 1743  
Oy 1921 TAGGCTCGGCGGCTGACGAGCATCAAAATATCGACGCTCAAGTCAAGAGTGGCGGAA 1980  
Db 1742 TAGGCTCGGCGGCTGACGAGCATCAAAATATCGACGCTCAAGTCAAGAGTGGCGGAA 1683  
Oy 1981 CCGGACAGGACTAATAAGATACAGAGGCTTCCCTGGAAGCTCCCTGTCGCTCTCC 2040  
Db 1682 CCGGACAGGACTAATAAGATACAGAGGCTTCCCTGGAAGCTCCCTGTCGCTCTCC 1623  
Oy 2041 TGTTCGACCTGCGCTTACCGGATACCTGTGCGCTTTCTCCCTTGGGAAAGCTGGC 2100  
Db 1622 TGTTCGACCTGCGCTTACCGGATACCTGTGCGCTTTCTCCCTTGGGAAAGCTGGC 1563  
Oy 2101 GCTTCTCATAGCTCAAGCTGTATGTAATCTCAGTTCGCTGATAGGTCCGTTGCTCAAGCT 2160  
Db 1562 GCTTCTCATAGCTCAAGCTGTATGTAATCTCAGTTCGCTGATAGGTCCGTTGCTCAAGCT 1503  
Oy 2161 GGGCTGTGTGACGAACCCCGCTTCAAGCCGACCGCTGGGCTTATTCGGTAACTATCG 2220  
Db 1502 GGGCTGTGTGACGAACCCCGCTTCAAGCCGACCGCTGGGCTTATTCGGTAACTATCG 1443  
Oy 2221 TCTTGAATCCACCCGTTAAGACAGCACTTATGCGCACTGGCAGACCACTGTTAAAG 2280  
Db 1442 TCTTGAATCCACCCGTTAAGACAGCACTTATGCGCACTGGCAGACCACTGTTAAAG 1383  
Oy 2281 GATTAGAGAGCGAGTATGATGAGCGGTGCTACAGATTTCTTGAAGTGGTGGCTAACTA 2340  
Db 1382 GATTAGAGAGCGAGTATGATGAGCGGTGCTACAGATTTCTTGAAGTGGTGGCTAACTA 1323  
Oy 2341 CCGCTACACTAAGAGCAAGTATTTGTATCTGCGCTCTGCTGAAGCAAGTAACTTCTCG 2400  
Db 1322 CCGCTACACTAAGAGCAAGTATTTGTATCTGCGCTCTGCTGAAGCAAGTAACTTCTCG 1263

```

QY 2401 AAAAAGTGTGAGCTCTTGAATCCGCAAAACCAACCGCTGTAGCGGTGTTTTT 2460
DB 1262 AAAAAGTGTGAGCTCTTGAATCCGCAAAACCAACCGCTGTAGCGGTGTTTTT 1203
QY 2461 TGTTCGACAGCAGATTAACGCGCAAAAAAGATCTCAAGAAATCTTTGATCTT 2520
DB 1202 TGTTCGACAGCAGATTAACGCGCAAAAAAGATCTCAAGAAATCTTTGATCTT 1143
QY 2521 TTTACGGGGTCTGACGCTCAGTGAACGAAAACCTCAGTTAAGGATTTTGTATAG 2580
DB 1142 TTTACGGGGTCTGACGCTCAGTGAACGAAAACCTCAGTTAAGGATTTTGTATAG 1083
QY 2581 ATTATCAAAAAAGATCTTCACTAGATCTTTTAATTAATAAGTTTAAATCAAT 2640
DB 1082 ATTATCAAAAAAGATCTTCACTAGATCTTTTAATTAATAAGTTTAAATCAAT 1023
QY 2641 CTAAAGTATATATAGATTAACCTGTCTGACAGTTACCAATGCTTAATCAGTGAAGCACC 963
DB 1022 CTAAAGTATATATAGATTAACCTGTCTGACAGTTACCAATGCTTAATCAGTGAAGCACC 963
QY 2701 TATCTAGGATCTGTCTATTTTCGTTCAATCAATAGTTCCTGCTGCTGATAT 2760
DB 962 TATCTAGGATCTGTCTATTTTCGTTCAATCAATAGTTCCTGCTGCTGATAT 2760
QY 2761 AACTAGATACGGAGGGCTTACCATCTGCGCCCAAGTGTCAATATACCGAGAACCC 2820
DB 902 AACTAGATACGGAGGGCTTACCATCTGCGCCCAAGTGTCAATATACCGAGAACCC 2820
QY 2821 ACGCTACGGGCTCAGATTTATCAGCAATTAACAGCCAGCGGAAAGGCGGAGCGCAG 2880
DB 842 ACGCTACGGGCTCAGATTTATCAGCAATTAACAGCCAGCGGAAAGGCGGAGCGCAG 2880
QY 2881 AAGTGTCTGTCACTTATCCGCTCCATCCAGTCTATTAATTTGTCGGGAAAGCTAG 2940
DB 782 AAGTGTCTGTCACTTATCCGCTCCATCCAGTCTATTAATTTGTCGGGAAAGCTAG 723
QY 2941 AGTAAGTATGTCGCACTTATATGTTTGGCAACGTTGTGCAATGCTAAGCAATCGT 3000
DB 722 AGTAAGTATGTCGCACTTATATGTTTGGCAACGTTGTGCAATGCTAAGCAATCGT 3000
QY 3001 GGTGTCAAGCTGTCTGTTGTATGAGCTTCAATGAGCTCCGCTTCCCAAGATCAAGGCG 603
DB 662 GGTGTCAAGCTGTCTGTTGTATGAGCTTCAATGAGCTCCGCTTCCCAAGATCAAGGCG 603
QY 3061 AGTTACATGATCCCAATGTTGTGCAAAAAAGCGTTAGCTCTTGCGTCTCCGATCGT 3120
DB 602 AGTTACATGATCCCAATGTTGTGCAAAAAAGCGTTAGCTCTTGCGTCTCCGATCGT 3120
QY 3121 TGTCAAGATTAAGTGGCGCAGATGTTATCATCATAGGTTATGCGACCTGCATATTC 483
DB 542 TGTCAAGATTAAGTGGCGCAGATGTTATCATCATAGGTTATGCGACCTGCATATTC 483
QY 3181 TCTTACTGTCAATGCGCATCCGTAAGATGCTTTCTGTGACTGTGTGATCTCAACCAATGC 3240
DB 482 TCTTACTGTCAATGCGCATCCGTAAGATGCTTTCTGTGACTGTGTGATCTCAACCAATGC 423
QY 3241 ATTCTGGAATTAAGTATGCGGAGCAGCGAGTGTCTTGGCCGCGGTCAATAAGGATTA 3300
DB 422 ATTCTGGAATTAAGTATGCGGAGCAGCGAGTGTCTTGGCCGCGGTCAATAAGGATTA 3300
QY 3301 TACCGGCGCAGATAGCAGAACTTAAAGTCTCATCATATTGGAAGAGTTCTTCCGAGGCG 3360
DB 362 TACCGGCGCAGATAGCAGAACTTAAAGTCTCATCATATTGGAAGAGTTCTTCCGAGGCG 3360
QY 3361 AAAAATCTCAAGATTTTACCGCTGTGAGATCCAGTTTCAATGTAACCCCTGTGTACCC 3420
DB 302 AAAAATCTCAAGATTTTACCGCTGTGAGATCCAGTTTCAATGTAACCCCTGTGTACCC 3420
QY 3421 CAATGATCTTCAAGATCTTTTACCGCTGTGAGATCCAGTTTCAATGTAACCCCTGTGTACCC 243
DB 242 CAATGATCTTCAAGATCTTTTACCGCTGTGAGATCCAGTTTCAATGTAACCCCTGTGTACCC 3480
QY 3481 GCAAAATGCCCAAAAAAGGAAATTAAGGCGACACGGAATGTTGAATCTCATCTCTT 3540

```

```

DB 182 GCAAAATGCGCAAAAAAGGAAATTAAGGCGACACGGAATGTTGAATCTCATCTCTT 123
QY 3541 CCTTTTCAATATTAATTAAGACATTTTATCAGGTTATTTGCTCATGAGGAGATTAAT 3600
DB 122 CCTTTTCAATATTAATTAAGACATTTTATCAGGTTATTTGCTCATGAGGAGATTAAT 63
QY 3601 TGAATGTTTATTAAGAAAAATTAACAAATAGGAGTTTCCGCGACATTTCCCGAAAGTCC 3660
DB 62 TGAATGTTTATTAAGAAAAATTAACAAATAGGAGTTTCCGCGACATTTCCCGAAAGTCC 3
QY 3661 AC 3662
DB 2 AC 1

RESULT 15
US-10-066-390-3/c
; Sequence 3, Application US/10066390
; Publication No. US20020146732A1
; GENERAL INFORMATION:
; APPLICANT: Padgett, Hal S.
; APPLICANT: Lindbo, John A.
; TITLE OF INVENTION: A Method of Increasing Complementarity
; FILE REFERENCE: P-UG 4878
; CURRENT APPLICATION NUMBER: US/10/066,390
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; LENGTH: 3637
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic construct
US-10-066-390-3

Query Match      72.5%; Score 2656.4; DB 13; Length 3637;
Best Local Similarity 84.0%; Pred. No. 0;
Matches 3079; Conservative 0; Mismatches 556; Indels 29; Gaps 6;

QY 1 CTAAATGTGAACGCTTAATATTTGTTAAATGCGCTTAATTTTGTAAATCAAGCTC 60
DB 3637 CTAAATGTGAACGCTTAATATTTGTTAAATGCGCTTAATTTTGTAAATCAAGCTC 3578
QY 61 ATTTTAAACCAATAGGCGCAAAATCCGCAAAATCCCTTAATTAATCAAAAGATAGCCGA 120
DB 3577 ATTTTAAACCAATAGGCGCAAAATCCGCAAAATCCCTTAATTAATCAAAAGATAGCCGA 3518
QY 121 GATTAGGTTGAGTGTGTTCCAGTTTGAACAGAGTCCCATTAATTAAGAAAGTGAAGT 180
DB 3517 GATTAGGTTGAGTGTGTTCCAGTTTGAACAGAGTCCCATTAATTAAGAAAGTGAAGT 3458
QY 3517 GATTAGGTTGAGTGTGTTCCAGTTTGAACAGAGTCCCATTAATTAAGAAAGTGAAGT 3458
DB 181 CAAGGTCAAAAGGCGGAAACCGTCTATCAAGGCGATGCGCCCATTAAGTGAACCAATC 240
QY 3457 CAAGGTCAAAAGGCGGAAACCGTCTATCAAGGCGATGCGCCCATTAAGTGAACCAATC 3398
DB 241 CTATCAAGTTTATTTGAGGTTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 300
QY 3397 CTATCAAGTTTATTTTGGGTTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 3338
DB 301 CCCCAGATTAGAGCTTGAACGGGAAACCGCGCAACGTGCGAGAAAGGAAAGGAAAGAA 360
QY 3337 CCCCAGATTAGAGCTTGAACGGGAAACCGCGCAACGTGCGAGAAAGGAAAGGAAAGAA 3278
DB 361 AGCGAAAGAGGCGCGCTTGAAGGCGCTGCAAGTGTAGCGGTCACCTGCGGTAAACAC 420
QY 3277 AGCGAAAGAGGCGCGCTTGAAGGCGCTGCAAGTGTAGCGGTCACCTGCGGTAAACAC 3218
DB 421 CACACCGCGCGCTTATAGCGCGCTACAGGCGCGGTCCCATTCGCAATCAGGCTGCG 480

```

Db 3217 CACACCGCGCGCTTAATGCGCGCTACAGGGCGCTCCATTGCCATTCAAGGCTCG 3158  
Qy 481 CAACTTTGGGAAGGCGATCGGTGGGCGCTTTCCTATTAGCCAGCTGGCGAAG 540  
Db 3157 CAACGTGTGGGAAGGCGATCGGTGGGCGCTTTCCTATTAGCCAGCTGGCGAAG 3098  
Qy 541 GGGATGTGCTCAAGGCGATTAGTTGGTTAAACGACAGGGTTTCCAGTACAGCGTTG 600  
Db 3097 GGGATGTGCTCAAGGCGATTAGTTGGTTAAACGACAGGGTTTCCAGTACAGCGTTG 3038  
Qy 601 TAAACGACGCGCAGTGAAGCGCGCTCGTTCAATTCACGTTTGTAAACCGTGAAGAG 660  
Db 3037 TAAACGACGCGCAGTGAAGCGCGGTATACGACTCATATAGGCGAATTGAGCTCCA 2978  
Qy 661 GCACATCGCGGTGCAAAATGTGTTTACAGCGTGAAGAGATGAAGATGCTGACAC 720  
Db 2977 CCGCGGTGGCGCGCTCAAACTAGTGAATCCCGCGCGTGAAGAAATTTATTTGT 2918  
Qy 721 GCTGCAAGACGCGAGCTAGATTAACTTAGAAGATATATATGTAAGTACGTTAA 780  
Db 2917 ATATGTTCAATCCATGCTATGTATATCCAGCGCTGTTACAAACTCAAGAGACATGT 2858  
Qy 781 AGATATATC-ATGCGTAAATTTGACGATGGGATCTGTAATACGACTCATATAGGCGA 838  
Db 2857 GGTCTCTTTTTCGTGGGATCTTTCGAAAGGCGAGATGTGTGACAGTAAATGTGT 2798  
Qy 839 ATTGGGTACCGGGCCCCCTCGAGGTCGACGTAATGATAGCTTGATATGCAATTCCT 898  
Db 2797 CTGTAAAGACAGAGGCTATGCGCAATTGAGATATTTGTTGATATATGCTGATGTT 2738  
Qy 899 GCACCGCGGGGATTCCTAGTTCTAGAGCGCGCGCACCGCGTGAAGCTCGAGCTTT 958  
Db 2737 GAAAGCTTCAATCTTCAATGTGTGTCTAATTTTGAAGTTAACTTTGATTCATTTT 2678  
Qy 959 GTTCCCTTATGAGGTTAATTAGATCCATGCGCTAATTTTACGACATATCTTCT 1018  
Db 2677 GTTGTGTGCAATGATGTAATATTGTGTGA-GTTATAGTTGTATTCAAATTTGTGTCA 2619  
Qy 1019 AGGGTAACTAGCTGATCAGATCATATGTCGGGCTTTTTCGCGCTCAGTATCG 1078  
Db 2618 AGAATGTTTCATCTTCTTAAATCAATACCTTTTAACTGATTC-----TATTAACA 2564  
Qy 1079 CCCAAGCTGGGCTATCTGGGCTATCGGGCATCGGGAGAAAGACCGCTGCTTTCCCGAGGT 1138  
Db 2563 GGGTATACCTTCAAACTTGACTTC-----AGCAGTGTCTTGTAGTTCCCTC 2515  
Qy 1139 TGAAGCGCATGAAAGAGTTTGCAGAGATGACTGCTGATGACTGATGAGCGAA 1198  
Db 2514 ATCTTGAATAATATAGTTCTTCTGTACATTAACCTTGGGATGACATCTTGAAATA 2455  
Qy 1199 ACGCAGCTTAACTATGATATTCGGGAAGTGTGGCCATGACGCTTTAAACGTTAACT 1258  
Db 2454 GTCATGCGCTTCAATGATCTGGGTATCTTGAAGAGATGAACACCATTAAGAAAGT 2395  
Qy 1259 GTTGTTCAGGCGCATGCGGATACCAATTCGTCGGGCTTTTCCGGAACAGTTCCGAT 1318  
Db 2394 AGTACAAGTGTGGCCATGGAACAGTATTTTCAAGTATGCAATAATTAATTAAGGT 2335  
Qy 1319 GGTGAGCCCGAAGCGCATCAGCAACCGAACAATACGCGGACAGCGGAACTGCGGTG 1378  
Db 2334 AAGTTTTC-----GTAATGTGCAATCACTTCACTTCCATGACGAAATTTGTGC 2280  
Qy 1379 CGGTGTGCAATTAATGACAGCGGTGCGCGCTGGGATATTACGTACAGCAGAGAGGTA 1438  
Db 2279 CCAATTAACATCAATCTAATTTCAACAAGAAATTTGGGCAACCTCCAGTAAAGTTCTTCT 2220  
Qy 1439 TCCGTGCTGATGCGCAGAAATGGAATGATACCCGTGATTTAACCCGCGCGCGCGC 1498  
Db 2219 CTTTACTCATCGGTACCACTTTTGTCTTTTATGAGAGGTTA-----ATTGCGGC 2165  
Qy 1499 TTGGCGTAACTGATGATGCTGTTTCTGTGTGAATTTGTAATCCGCTCACAATTTCA 1558  
Db 2164 TTGGCGTAACTGATGATGCTGTTTCTGTGTGAATTTGTAATCCGCTCACAATTTCA 2105

Qy 1559 CACAACTATGAGCGCGGAAGATTAAGTTAAACCTGGGGTGCCTAATGATGAGCTTA 1618  
Db 2104 CACAACTATGAGCGCGGAAGATTAAGTTAAACCTGGGGTGCCTAATGATGAGCTTA 2045  
Qy 1619 CTCACTAATTAATGCGGTGCGCTCACTGCGCGCTTTCAGTGGGAAACCTGTCGTCAG 1678  
Db 2044 CTCACTAATTAATGCGGTGCGCTCACTGCGCGCTTTCAGTGGGAAACCTGTCGTCAG 1985  
Qy 1679 CTGCAATTAATGAGTGGCCAAACGCGCGGGAGAGCGGTTTGCCTAATTTGGCGCTTTCC 1738  
Db 1984 CTGCAATTAATGAGTGGCCAAACGCGCGGGAGAGCGGTTTGCCTAATTTGGCGCTTTCC 1925  
Qy 1739 GCTTCTGCTCACTGACTGCTGCTCGCTGCTTGTGGCTGGCGGACGCTATCACT 1798  
Db 1924 GCTTCTGCTCACTGACTGCTGCTGCTGCTGCTTGTGGCTGGCGGACGCTATCACT 1865  
Qy 1799 CACTCAAGGCGGTATACGGTTATCCAGAAATCAGGGGATTAACGAGAAAGAACATG 1858  
Db 1864 CACTCAAGGCGGTATACGGTTATCCAGAAATCAGGGGATTAACGAGAAAGAACATG 1805  
Qy 1859 TGACCAAAAGGCGCAAAAGGCGGAAACGTTAAAGGCGGTTGCTGCTGCTTTTTC 1918  
Db 1804 TGACCAAAAGGCGCAAAAGGCGGAAACGTTAAAGGCGGTTGCTGCTGCTTTTTC 1745  
Qy 1919 CATAGGCTCGCGCGCGCTGACAGATCACAAAATGACGCTCAAGTCAAGTGGCGCA 1978  
Db 1744 CATAGGCTCGCGCGCGCTGACAGATCACAAAATGACGCTCAAGTCAAGTGGCGCA 1685  
Qy 1979 AACCGACAGGACTATTAAGTATACAGGCGTTTCCCTGGAACCTCCTGTGCGCTCT 2038  
Db 1684 AACCGACAGGACTATTAAGTATACAGGCGTTTCCCTGGAACCTCCTGTGCGCTCT 1625  
Qy 2039 CTTGTTCCGACCCCTGCGCTTACCGGATACCTGTCGCGCTTTCCTTTCGGAAGCGTG 2098  
Db 1524 CTTGTTCCGACCCCTGCGCTTACCGGATACCTGTCGCGCTTTCCTTTCGGAAGCGTG 1565  
Qy 2099 GCGCTTCTCATAGCTCACTGATGATGATCTCACTTGTGATGATGCTGCTCAAG 2158  
Db 1564 GCGCTTCTCATAGCTCACTGATGATGATCTCACTTGTGATGATGCTGCTCAAG 1505  
Qy 2159 CTGGGCTGTGACAGAAACCCCGCTTACGCGGACCGCTGCGCTTATCCGGTAACAT 2218  
Db 1504 CTGGGCTGTGACAGAAACCCCGCTTACGCGGACCGCTGCGCTTATCCGGTAACAT 1445  
Qy 2219 CGTCTTGAATCAACCCGTTAAGACAGACTTATGCGCACTGGACGACCACTGTATAC 2278  
Db 1444 CGTCTTGAATCAACCCGTTAAGACAGACTTATGCGCACTGGACGACCACTGTATAC 1385  
Qy 2279 AGAATTAAGCAGAGGATATGAGCGGTGCTACAGAGTTCTTGAAGTGTGGCTTAAC 2338  
Db 1384 AGAATTAAGCAGAGGATATGAGCGGTGCTACAGAGTTCTTGAAGTGTGGCTTAAC 1335  
Qy 2339 TAGGCTTACCTAAGAAAGATTTGTATGCTGCGCTGCTGGAAGCCAGTAACTTC 2398  
Db 1324 TAGGCTTACCTAAGAAAGATTTGTATGCTGCGCTGCTGGAAGCCAGTAACTTC 1265  
Qy 2399 GGAAGAAAGATGTTGTTACTCTTGTATCCGGCAACAAACACCGCTGATGCGGTGTTT 2458  
Db 1264 GGAAGAAAGATGTTGTTACTCTTGTATCCGGCAACAAACACCGCTGATGCGGTGTTT 1205  
Qy 2459 TTTGTTTGAAGCAGCAGATTAACGCGGAAGAAAGAGATCTCAAGAGATCTTTGATC 2518  
Db 1204 TTTGTTTGAAGCAGCAGATTAACGCGGAAGAAAGAGATCTCAAGAGATCTTTGATC 1145  
Qy 2519 TTTTCTAAGGAGTCTGACGCTCACTGGAAGAAACCTCAAGTAAAGGATTTTGTATG 2578  
Db 1144 TTTTCTAAGGAGTCTGACGCTCACTGGAAGAAACCTCAAGTAAAGGATTTTGTATG 1085  
Qy 2579 AGATTATCAAAAGAGATCTTCACTTATGATCTTTTAAATTAATTAAGTTTAAATCA 2638  
Db 1084 AGATTATCAAAAGAGATCTTCACTTATGATCTTTTAAATTAATTAAGTTTAAATCA 1025

QY 2639 ATCTAAGTATATATAGTAAACTGTGTCAGACATTACCATCTTAAATCAGTAGCA 2698  
Db 1024 ATCTAAGTATATATAGTAACTGTGTCAGACATTACCATCTTAAATCAGTAGCA 965  
QY 2699 CCTATCTCAGCGATCTGTCTATTTTCGTTTCATCTCATAGTGGCTGACCTCCGTCGTGTAG 2758  
Db 964 CCTATCTCAGCGATCTGTCTATTTTCGTTTCATCTCATAGTGGCTGACCTCCGTCGTGTAG 905  
QY 2759 ATTAAGTATAGTATAGTAAAGGAGGCTTACCATCTGCTGACATGATATCCGAGAC 2818  
Db 904 ATTAAGTATAGTAAAGGAGGCTTACCATCTGCTGACATGATATCCGAGAC 845  
QY 2819 CCAAGCTTCAACGGCTTCAAGATTATCAGCAATAAACCAAGCCAGCCGGAAGGCGGAGCC 2878  
Db 844 CCAAGCTTCAACGGCTTCAAGATTATCAGCAATAAACCAAGCCAGCCGGAAGGCGGAGCC 785  
QY 2879 AGAAGTGTCTCTGAATTTATCCGCTCCATCCAGTCTATTAATTGTCGAGAAAGCT 2938  
Db 784 AGAAGTGTCTCTGAATTTATCCGCTCCATCCAGTCTATTAATTGTCGAGAAAGCT 725  
QY 2939 AGAAGTATAGTATCGCCAGTTAAATGTTGCGCAAGTGTGTCATGCTTCCAGCATC 2998  
Db 724 AGAAGTATAGTATCGCCAGTTAAATGTTGCGCAAGTGTGTCATGCTTCCAGCATC 665  
QY 2999 GTGGTTCAGCGCTCGCTGTTGTTGATGCTTCAATCAAGCTCCGTTCCCAAGCATCAAG 3058  
Db 664 GTGGTTCAGCGCTCGCTGTTGTTGATGCTTCAATCAAGCTCCGTTCCCAAGCATCAAG 605  
QY 3059 CGAGTTACATGATCCCCCATGTTGTCAAAAAAGCGTTAGCTCTTGGTCCGATC 3118  
Db 604 CGAGTTACATGATCCCCCATGTTGTCAAAAAAGCGTTAGCTCTTGGTCCGATC 545  
QY 3119 GTTGTCAAGTATAGTGGCCGCACTGTTATCATCTCATGTTATGCGACACTGCATAT 3178  
Db 544 GTTGTCAAGTATAGTGGCCGCACTGTTATCATCTCATGTTATGCGACACTGCATAT 485  
QY 3179 TCTTTACTGTATGCAATCCGTAGATGCTTTCTGATCTGTCAGTACTCAACCAAG 3238  
Db 484 TCTTTACTGTATGCAATCCGTAGATGCTTTCTGATCTGTCAGTACTCAACCAAG 425  
QY 3239 TCAATCTGAGAAATGATGAGGCGGACCGAGTGTCTTGGCCGCGTCAATACGGGAT 3298  
Db 424 TCAATCTGAGAAATGATGAGGCGGACCGAGTGTCTTGGCCGCGTCAATACGGGAT 365  
QY 3299 AATACCGGCGCACATAGCAGAACTTTAAAGTGTCTCATTTGAAACGTTCTTCGGAG 3358  
Db 364 AATACCGGCGCACATAGCAGAACTTTAAAGTGTCTCATTTGAAACGTTCTTCGGAG 305  
QY 3359 CGAAACTCTCAAGATTTTACGCTGTGATCCAGTTCAGTATCCCACTCGTGA 3418  
Db 304 CGAAACTCTCAAGATTTTACGCTGTGATCCAGTTCAGTATCCCACTCGTGA 245  
QY 3419 CCAACTGATCTTCAAGATTTTACTTTCACCAAGGCTTCTGGTGAACAAACAGGA 3478  
Db 244 CCAACTGATCTTCAAGATTTTACTTTCACCAAGGCTTCTGGTGAACAAACAGGA 185  
QY 3479 AGGCAAAATGCGCAAAAGGGAATAGGCGCACAGGAAATGTTGAATCTCATCTC 3538  
Db 184 AGGCAAAATGCGCAAAAGGGAATAGGCGCACAGGAAATGTTGAATCTCATCTC 125  
QY 3539 TTCTTTTCAATATTTATGAGCATTTATCAGGGTTATTTGTCATGAGCGGATACATA 3598  
Db 124 TTCTTTTCAATATTTATGAGCATTTATCAGGGTTATTTGTCATGAGCGGATACATA 65  
QY 3599 TTTGAATGTTATTTAGAAAAATTAACAAATAGGGGTTCCGCGACATTTCCCGAAAAAGTG 3658  
Db 64 TTTGAATGTTATTTAGAAAAATTAACAAATAGGGGTTCCGCGACATTTCCCGAAAAAGTG 5  
QY 3659 CCAC 3662  
Db 4 CCAC 1

Search completed: January 26, 2005, 17:51:53  
Job time : 1991.66 sec8

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

## OM nucleic - nucleic search, using sw model

Run on: January 26, 2005, 02:19:36 ; Search time 304.259 Seconds

(without alignments)  
8554.922 Million cell updates/sec

Title: US-10-826-523-41

Perfect score: 3662  
Sequence: 1 ctaaatcgttaagcgttaata.....attccccgaagtcac 3662

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq:\*  
2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq:\*  
3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq:\*  
4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq:\*  
5: /cgn2\_6/ptodata/1/ina/PTUS\_COMB.seq:\*  
6: /cgn2\_6/ptodata/1/ina/backfile1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2639	72.1	3681	1 US-07-924-028A-6	Sequence 6, Appli
2	2623	71.6	5314	1 US-07-924-028A-4	Sequence 4, Appli
3	2620	71.6	3699	3 US-08-646-538-6	Sequence 6, Appli
4	2620	71.6	3699	3 US-09-503-222-6	Sequence 6, Appli
5	2469	67.4	5277	4 US-09-890-848-16	Sequence 16, Appli
6	2448	66.9	3956	4 US-09-402-266B-21	Sequence 21, Appli
7	2305	62.9	4102	4 US-09-402-266B-20	Sequence 20, Appli
8	2302	62.9	4088	4 US-09-402-266B-18	Sequence 18, Appli
9	2279	62.2	4119	4 US-09-993-170-1	Sequence 1, Appli
10	2172	59.3	7102	3 US-09-138-024-20	Sequence 20, Appli
11	2172	59.3	7102	3 US-09-404-066-20	Sequence 20, Appli
12	2172	59.3	7102	4 US-09-573-322-20	Sequence 20, Appli
13	2172	59.3	7333	4 US-09-138-024-21	Sequence 21, Appli
14	2172	59.3	7333	4 US-09-404-066-21	Sequence 21, Appli
15	2172	59.3	7333	4 US-09-573-322-21	Sequence 21, Appli
16	2170	59.3	5178	2 US-08-474-169-2	Sequence 2, Appli
17	2170	59.3	6206	2 US-08-474-169-3	Sequence 3, Appli
18	2170	59.3	2973	4 US-09-402-266B-17	Sequence 17, Appli
19	2170	59.3	3789	4 US-09-075-019-8	Sequence 8, Appli
20	2170	59.3	4583	4 US-09-402-266B-19	Sequence 19, Appli
21	2170	59.3	2961	3 US-08-446-935-6	Sequence 6, Appli
22	2170	59.3	4145	3 US-08-651-472-62	Sequence 62, Appli
23	2170	59.3	4145	3 US-08-358-928-62	Sequence 62, Appli
24	2170	59.3	4277	3 US-08-651-472-63	Sequence 63, Appli
25	2170	59.3	4277	3 US-08-358-928-63	Sequence 63, Appli
26	2170	59.3	5356	3 US-08-446-935-1	Sequence 1, Appli
27	2170	59.3	5532	3 US-08-651-472-72	Sequence 72, Appli

C 28	2170	59.3	5532	3 US-08-358-928-72	Sequence 72, Appli
C 29	2170	59.3	5534	1 US-08-452-267-3	Sequence 3, Appli
C 30	2170	59.3	5534	3 US-09-123-644-3	Sequence 3, Appli
C 31	2170	59.3	6811	3 US-08-651-472-67	Sequence 67, Appli
C 32	2170	59.3	6811	3 US-08-358-928-67	Sequence 67, Appli
C 33	2170	59.3	6926	3 US-08-651-472-69	Sequence 69, Appli
C 34	2170	59.3	6926	3 US-08-358-928-69	Sequence 69, Appli
C 35	2170	59.3	7090	4 US-09-714-520-18	Sequence 18, Appli
C 36	2170	59.3	7214	4 US-09-438-268-1	Sequence 1, Appli
C 37	2170	59.3	7432	4 US-09-367-851A-6	Sequence 6, Appli
C 38	2170	59.3	7474	4 US-09-845-917A-26	Sequence 26, Appli
C 39	2170	59.3	8560	4 US-09-936-572-11	Sequence 11, Appli
C 40	2170	59.3	11233	3 US-08-980-832-27	Sequence 27, Appli
C 41	2170	59.3	11233	4 US-09-920-923B-27	Sequence 27, Appli
C 42	2169	59.2	3737	2 US-08-784-208-1	Sequence 1, Appli
C 43	2169	59.2	3796	4 US-09-470-661A-32	Sequence 32, Appli
C 44	2169	59.2	4399	4 US-09-911-781-27	Sequence 27, Appli
C 45	2169	59.2	4399	4 US-10-400-902-27	Sequence 27, Appli

## ALIGNMENTS

RESULT 1

US-07-924-028A-6  
Sequence 6, Application US/07924028A

Patent No. 5470573

GENERAL INFORMATION:

APPLICANT: Lubitz Werner, Szoestak, Michael P.

TITLE OF INVENTION: CARRIER-BOUND RECOMBINANT PROTEINS, PROCESS

FOR THE PRODUCTION AND USE AS IMMUNOGENS AND VACCINES

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:

ADDRESSEE: Felfe &amp; Lynch

STREET: 805 Third Avenue

CITY: New York City

STATE: New York

COUNTRY: USA

ZIP: 10022

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage

COMPUTER: IBM PS/2

OPERATING SYSTEM: PC-DOS

SOFTWARE: Wordperfect

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/924,028A

FILING DATE: 30-SEP-1992

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/EP91/00308

FILING DATE: 02-FEB-1991

PRIOR APPLICATION DATA:

APPLICATION NUMBER: DE 40 05 874

FILING DATE: 24-FEB-1990

ATTORNEY/AGENT INFORMATION:

NAME: Hanson, No. 5470573man D.

REGISTRATION NUMBER: 30,946

REFERENCE/DOCKET NUMBER: HUBR 1027

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 688-9200

TELEFAX: (212) 838-3884

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 3681 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

US-07-924-028A-6

Query Match 72.1%, Score 2639, DB 1, Length 3681;  
Best Local Similarity 83.5%, Pred. No. 0;  
Matches 3075, Conservative 0, Mismatches 580, Indels 28, Gaps 6;

QY 3 AAATGTAAGCGTTAATATTTTGTAAATTCGCGTTAAATTTTGTAAATCACTCAT 62  
Db 1 AAATTTTAAACGTATATATTTTGTAAATTCGCGTTAAATTTTGTAAATCACTCAT 60  
QY 63 TTTTAAACCAATAGCGCGAAATTCGCGAAATTCCTTAAATCAAAAGATAGACCGAGA 122  
Db 61 TTTTAAACCAATAGCGCGAAATTCGCGAAATTCCTTAAATCAAAAGATAGACCGAGA 120  
QY 123 TAGGGTTGAGTGTGTTCCAGTTTGGAAACAAGTCCACTATTAAGAAAGTGGACTCCA 182  
Db 121 TAGGGTTGAGTGTGTTCCAGTTTGGAAACAAGTCCACTATTAAGAAAGTGGACTCCA 180  
QY 183 AOCGTAAAGGCGGAAAAACCGTCTATCAGGAGGATGGCCCACTAGCTGAACCATCCTT 242  
Db 181 ACGTAAAGGCGGAAAAACCGTCTATCAGGAGGATGGCCCACTAGCTGAACCATCCTT 240  
QY 243 AATCAAGTTTGTGGGGTGAAGGTGCCCTTAAGCACTAAATCGAAACCTTAAGGAGCC 302  
Db 241 AATCAAGTTTGTGGGGTGAAGGTGCCCTTAAGCACTAAATCGAAACCTTAAGGAGCC 300  
QY 303 CCGAATTAGACTTGAACGAGGAAAGCCGCGCAAGCTGGCGAAGAAAGGAAAGAAAG 362  
Db 301 CCGAATTAGACTTGAACGAGGAAAGCCGCGCAAGCTGGCGAAGAAAGGAAAGAAAG 360  
QY 363 CGAAAGAGCGGCGCGCTAGGAGCGCTGAGCAAGTGTAGCGGTCAAGCTGGCGAAGAAAG 422  
Db 361 CGAAAGAGCGGCGCGCTAGGAGCGCTGAGCAAGTGTAGCGGTCAAGCTGGCGAAGAAAG 420  
QY 423 CACCGCGCGCGCTTATGCGCGCGCTACAGGCGCGCTCCATTTGCGCAATCAGGCTCGCA 482  
Db 421 CACCGCGCGCGCTTATGCGCGCGCTACAGGCGCGCTCCATTTGCGCAATCAGGCTCGCA 480  
QY 483 ACTGTGGGAAAGGCGCATCGGTGCGGCGCTTCCGTATTAAGCGCACTGGCGAAAGGGG 542  
Db 481 ACTGTGGGAAAGGCGCATCGGTGCGGCGCTTCCGTATTAAGCGCACTGGCGAAAGGGG 540  
QY 543 GATGTGCTGCAAGGCGCTTAAATTGGGTAAACGCAAGGGTTTCCAGTCAAGAGCTTTGA 602  
Db 541 GATGTGCTGCAAGGCGCTTAAATTGGGTAAACGCAAGGGTTTCCAGTCAAGAGCTTTGA 600  
QY 603 AAAGGACGCGCAATGAGCGCGCTCGTTCACTTCAAGTTTGAACCGCTGGAGACGGCG 662  
Db 601 AAAGGACGCGCAATGAGCGCGCTCGTTCACTTCAAGTTTGAACCGCTGGAGACGGCG 660  
QY 663 AGACTCGCGGTGCAATGCTTTTACAGCGTATGAGACAGATGAAGTCTGACACCG 722  
Db 661 GTCGCGCGCTTATGATGAGCACTCTCAGTCAATCTGCTGTATGCCGATAGTTA 720  
QY 723 TGCAGAACACGCACTAGATTAACCTAGAAAGATTAATCAATTGTAAGCTAGTTAAG 782  
Db 721 AGCAGATATATACATTCGCTATCGCTACGTGACTGGGTGATGGCTGCCGACACCC 780  
QY 783 ATATATCATGCGTAAATTAAGCATGGATCTGTAAATACACTATTAAGGGCAATTTG 842  
Db 781 GCCAACACCGCGTACCGCTCTGACGCGCTGTCTGCTCCGCGACTCCGCTTAACACA 840  
QY 843 GGTACCGGCGCGCGCTCTGAGGTGACGATTCGA-----TAACTTGAATTCGAATTC 896  
Db 841 AGCTGTACCGCTCTGAGGTGACGATTCGA-----TAACTTGAATTCGAATTC 896  
QY 897 CTGCAAGCGCGGAGATTCATAGTCTAGAGCGCGCGCAACCGCGGTGAGACTCCAGTT 956  
Db 901 CGGAGGCAATAGTGTGATGCTTTGTGAGCAATTCGTCCCTTAAGTAAGCAATTCCTG 960  
QY 957 TTGTCCCTTATAGAGGTTAATTAAGATCCCAT-----GCTCAATTTTAAGCAGAC 1009  
Db 961 TAAAGTGTCACTGTGCGGATCAACGCTTCCAGTACGACAGAAAGCAATGATGTGTA 1020  
QY 1010 TATCTTTTAGGGTTAATCTAGCTGATCAGATCAATATGTCGGGTCTTTTCCGGCT 1069  
Db 1021 TTTTCAGAGAAAGATTCGAGAGAAATATATATCAATAAAGTTGAACTTTTGTGTCT 1080  
QY 1070 CAGTATGCGCCAACTGCGCGCTATCTGGGCAATCGGAGGAGAAAGACCGCGCTTTTC 1129

Db 1081 TCGACATGGGTAAATCTCATGTTTGAATGCGCCCTAGAGGATCCGCGCAACCTTGCAATGCC 1140  
QY 1130 CCGCGAGGTTGAAGGCGCATGAGAAAGTTTCCGAGATGACTGCTGCTGCAATTAAGCT 1189  
Db 1141 TCGAGGTGCACTTAGAGATTCGCCGACGCTCGACGCAATTAATATGTTTCCGTAAT 1200  
QY 1190 TGAAGGAAAGCGACGTTTACATGATGATTCGAGGAAAGGTGGATGCTGACGCTTTAA 1249  
Db 1201 TGAAGGCTTCATGATGAGACAGGCGCTTT--GAATGTGACGAGATGAAATATTAAG 1258  
QY 1250 CCGTGAATGTTTCTGAGCGCACTCGGATACCAAGTTGCTGCGGCTTTTCCGAGACA 1309  
Db 1259 CAATGAGCGGACGAATTAATCTCAACAG--AGCAGAAAGGAGGATATCCCAAAAGT 1315  
QY 1310 GTTCGAGATGTCAGCCCGAAGCGCATCAGCAACCGCAATTAACCGCGCAG----- 1363  
Db 1316 CCAAGCTACATTAACGCAAGCGCTCAAGCGCACGACAGAGCGGTCACTAGCA 1375  
QY 1364 CCGAATGCTGCGCGCGGTGACATTAATGACAGCGGTGCGCGCTGGGATTTATAGT 1423  
Db 1376 TCCAACTTTGTTACTCTGACAGAAATGCAATCATCTTCGTTAAATCCAAACCGCAG 1435  
QY 1424 CAGCGAGACGGGTAT---CTGCTGATGATCCGCAAGAAATGAGACATGATATCCCGTG 1479  
Db 1436 AAGCTGAATGAAATTCGACTGAGGGGGGCGGTACCAAGCTTTTGTCCCTTTA 1495  
QY 1480 AGTTACCGGCGCGCGCGCTTGGGTAAATGATGATCACTAGCTTTTCTGTGAAATTG 1539  
Db 1496 GTGAGGGTTAAATTCGAGCTTGGGTAAATGATGATGATGCTGTTCTGTGAAATTG 1555  
QY 1540 TTAATCCGCTCAATTTCAACAACAATACAGCGCGAACAATAAGTGAAGCCTGGGG 1599  
Db 1556 TTAATCCGCTCAATTTCAACAACAATACAGCGCGAACAATAAGTGAAGCCTGGGG 1615  
QY 1600 TGCCTAATGATGACTAATCACTTAATGCGTTGGCTTACTGCGCGTTTCAAGTC 1659  
Db 1616 TGCCTAATGATGACTAATCACTTAATGCGTTGGCTTACTGCGCGTTTCAAGTC 1675  
QY 1660 GGGAAACCTGCTGACAGCTGCAATTAATGAATCGGCCCAAGCGCGGAGAGCGGTT 1719  
Db 1676 GGGAAACCTGCTGACAGCTGCAATTAATGAATCGGCCCAAGCGCGGAGAGCGGTT 1735  
QY 1720 GCGTATTTGGCGCTTTCGCTTCTCGCTCACTGACCTGCTGCTCGCTTTCGCT 1779  
Db 1736 GCGTATTTGGCGCTTTCGCTTCTCGCTCACTGACCTGCTGCTCGCTTTCGCT 1795  
QY 1780 GCGGCGAGCGTATGACTCACTAAAGCGGTAAATCGGTATTCAGAGAAATCAGGGGA 1839  
Db 1796 GCGGCGAGCGTATGACTCACTAAAGCGGTAAATCGGTATTCAGAGAAATCAGGGGA 1855  
QY 1840 TAAAGCAAGAAAGAAACATGTGAGCAAAAGCGCAAAAGGCGCAAGAACCTGTAAGAGC 1899  
Db 1856 TAAAGCAAGAAAGAAACATGTGAGCAAAAGCGCAAAAGGCGCAAGAACCTGTAAGAGC 1915  
QY 1900 CCGCTTGTGCGCTTTTCAATAGGCTCGCGCCCTCTGACGACATCAAAATTCGACG 1959  
Db 1916 CCGCTTGTGCGCTTTTCAATAGGCTCGCGCCCTCTGACGACATCAAAATTCGACG 1975  
QY 1960 CTCAAGTCAAGGTGGGAAACCGCAACAGACTATTAAGATTAACAGAGGTTTCCCGCTG 2019  
Db 1976 CTCAAGTCAAGGTGGGAAACCGCAACAGACTATTAAGATTAACAGAGGTTTCCCGCTG 2035  
QY 2020 AAGCTCTCTGCGCTCTCTGTTCCGACCTGCGCTTACCGGATACCTGTCGCTT 2079  
Db 2036 AAGCTCTCTGCGCTCTCTGTTCCGACCTGCGCTTACCGGATACCTGTCGCTT 2095  
QY 2080 TCTCCCTTGGGAAAGCGTGGGCTTTTCTCATAGCTCAAGCTGATGATTCGCTT 2139  
Db 2096 TCTCCCTTGGGAAAGCGTGGGCTTTTCTCATAGCTCAAGCTGATGATTCGCTT 2155  
QY 2140 GTAGGTGCTTCAAGCTGAGCTGTGTGACGAACCCCGTTCAAGCCGACGCTG 2199



Db 2156 GTAGTCGTTGCTTCAAGCTGAGGCTGTGTGACGAAACCCCGCTTACGCCGACCGCTG 2215  
Qy 2200 CGCTTATCCGGTAACCTATGCTTTGAGTCCAAACCCGGTAAGACACGACTTATGGCCACT 2259  
Db 2216 CGCTTATCCGGTAACCTATGCTTTGAGTCCAAACCCGGTAAGACACGACTTATGGCCACT 2275  
Qy 2260 GGCAGCAGCCACTGTTAAACAGATTAGCAGACGCGAGTATGTAAGCGGTGCTACAGAGTT 2319  
Db 2276 GGCAGCAGCCACTGTTAAACAGATTAGCAGACGCGAGTATGTAAGCGGTGCTACAGAGTT 2335  
Qy 2320 CTGTAAGTGTGTGCTTAACTAAGCTGCTAAGAGGACAGTATTTGGTATCTGCGCTCT 2379  
Db 2336 CTGTAAGTGTGTGCTTAACTAAGCTGCTAAGAGGACAGTATTTGGTATCTGCGCTCT 2395  
Qy 2380 GCTAAGCCACTTACCTTCCGAAAAAGATTGTAAGTCTTGAATCCGCAACCAACAC 2439  
Db 2396 GCTAAGCCACTTACCTTCCGAAAAAGATTGTAAGTCTTGAATCCGCAACCAACAC 2455  
Qy 2440 CGCTGTAGCGGTGTTTGTGTTGCAAGCAGAGATTAGCGCGAGAAAAAGAGATC 2499  
Db 2456 CGCTGTAGCGGTGTTTGTGTTGCAAGCAGAGATTAGCGCGAGAAAAAGAGATC 2515  
Qy 2500 TCAAGAGATCTTGTGATCTTTCTACCGGGTCTGACGCTGACAGTGAACGAAAACTCAG 2559  
Db 2516 TCAAGAGATCTTGTGATCTTTCTACCGGGTCTGACGCTGACAGTGAACGAAAACTCAG 2575  
Qy 2560 TTAAGGATTTTGTGATGATATATCAAAAAAGATCTTCACTAGATCCTTTAAATTA 2619  
Db 2576 TTAAGGATTTTGTGATGATATATCAAAAAAGATCTTCACTAGATCCTTTAAATTA 2635  
Qy 2620 AAAATGAAGTTTAAATCAATCTAAGATATATGAGTAAACCTTGTGACAGTTTCCA 2679  
Db 2636 AAAATGAAGTTTAAATCAATCTAAGATATATGAGTAAACCTTGTGACAGTTTCCA 2695  
Qy 2680 ATGCTTATCAGTAGAGGACCTATCTCAGCGATCTGTCTATTTGTTTCATCATAGTTGC 2739  
Db 2696 ATGCTTATCAGTAGAGGACCTATCTCAGCGATCTGTCTATTTGTTTCATCATAGTTGC 2755  
Qy 2740 CTGACTCCCGTCTGTGATGATATCTAGCATCGGAGGGCTTACCATCTGCCCCAAGTGC 2799  
Db 2756 CTGACTCCCGTCTGTGATGATATCTAGCATCGGAGGGCTTACCATCTGCCCCAAGTGC 2815  
Qy 2800 TGCAATGATACCGGAGACCCACGCTCACCGGGCTCCAGATTATACAGAAATAAACGACC 2859  
Db 2816 TGCAATGATACCGGAGACCCACGCTCACCGGGCTCCAGATTATACAGAAATAAACGACC 2875  
Qy 2860 AGCCGGAAGGGCCGAGCGCAGAAAGTGTCTGCAACTTATCCGCTCCATCCAGTCTAT 2919  
Db 2876 AGCCGGAAGGGCCGAGCGCAGAAAGTGTCTGCAACTTATCCGCTCCATCCAGTCTAT 2935  
Qy 2920 TAAATGTTGCGGGAAGCTAGAGTAAAGTTCGCCAGTTAAATGTTTGCAGAACGTTGT 2979  
Db 2936 TAAATGTTGCGGGAAGCTAGAGTAAAGTTCGCCAGTTAAATGTTTGCAGAACGTTGT 2995  
Qy 2980 TGCCATTGCTACAGAGCATCGTGTGTGACGCTCGTGGTTGGTATGCTTATCACTC 3039  
Db 2996 TGCCATTGCTACAGAGCATCGTGTGTGACGCTCGTGGTTGGTATGCTTATCACTC 3055  
Qy 3040 CGGTTCCCAAGCATCAAGCGAGTAACTATCCCATGTTTGGCAAAAAAGCGGTAG 3099  
Db 3056 CGGTTCCCAAGCATCAAGCGAGTAACTATCCCATGTTTGGCAAAAAAGCGGTAG 3115  
Qy 3100 CTCTTTGGTCTCTCGATCGTTGTCAAGAGTAAAGTTGCGCGAGTATCACTCATGT 3159  
Db 3116 CTCTTTGGTCTCTCGATCGTTGTCAAGAGTAAAGTTGCGCGAGTATCACTCATGT 3175  
Qy 3160 TATGGCAGCACTGCTAATTTCTTACGTGATGCGCATCCGTAGAGGCTTTTCTGTGAC 3219  
Db 3176 TATGGCAGCACTGCTAATTTCTTACGTGATGCGCATCCGTAGAGGCTTTTCTGTGAC 3235  
Qy 3220 TGGTGAATCACTCAACCAAGTCAATCTAGAGTATGTAAGCGGACGAGTGTCTCTTG 3279  
Db 3236 TGGTGAATCACTCAACCAAGTCAATCTAGAGTATGTAAGCGGACGAGTGTCTCTTG 3295

Qy 3280 CCGCGCTCAATAGGGATTAATACCGGCGCCCATAGCAGAACTTTAAAGTCTCATCAT 3339  
Db 3396 CCGCGCTCAATAGGGATTAATACCGGCGCCCATAGCAGAACTTTAAAGTCTCATCAT 3355  
Qy 3340 TGAAGAGCTTCTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTGAGATCCAGTTTC 3399  
Db 3356 TGAAGAGCTTCTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTGAGATCCAGTTTC 3415  
Qy 3400 GATGTAAACCACTGCTGACACCACTGATCTTTCAGATCTTTTACCTTTCACAGCTTTC 3459  
Db 3416 GATGTAAACCACTGCTGACACCACTGATCTTTCAGATCTTTTACCTTTCACAGCTTTC 3475  
Qy 3460 TGGGTGAGCAAAAAAGAGAGCAAAATGCCGCAAAAAAGGATTAAGGGGACACGGA 3519  
Db 3476 TGGGTGAGCAAAAAAGAGAGCAAAATGCCGCAAAAAAGGATTAAGGGGACACGGA 3535  
Qy 3520 ATGTTGAATCTGATCTCTCTTTTCAATAATTATGAAGCAATTAATCAGGGTTATGT 3579  
Db 3536 ATGTTGAATCTGATCTCTCTTTTCAATAATTATGAAGCAATTAATCAGGGTTATGT 3595  
Qy 3580 TCTCATGAGCGGATACATAATTTGAATGATTTAGAAAAATTAACAAATAGGGGTTCCG 3639  
Db 3596 TCTCATGAGCGGATACATAATTTGAATGATTTAGAAAAATTAACAAATAGGGGTTCCG 3655  
Qy 3640 CACATTTCCCGAAAAAGTCCAC 3662  
Db 3656 CACATTTCCCGAAAAAGTCCAC 3678

RESULT 2  
US-07-924-028A-4  
; Sequence 4, Application US/07924028A  
; Patent No. 5470573  
; GENERAL INFORMATION:  
; APPLICANT: Lubitz Werner, Szostak, Michael P.  
; TITLE OF INVENTION: CARRIER-BOUND RECOMBINANT PROTEINS, PROCESS  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pelke & Lynch  
; STREET: 805 Third Avenue  
; CITY: New York City  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10022  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: PC-DOS  
; SOFTWARE: Wordperfect  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/924,028A  
; FILING DATE: 30-FEB-1992  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/EP91/00308  
; FILING DATE: 02-FEB-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: DE 40 05 874  
; FILING DATE: 24-FEB-1990  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Hanson, No. 5470573man D.  
; REGISTRATION NUMBER: 30,946  
; REFERENCE/DOCKET NUMBER: HUBR 1027  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 688-9200  
; TELEFAX: (212) 838-3884  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 5314 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single

; TOPOLOGY: linear

US-07-924-028A-4

Query Match Best Local Similarity 71.6%; Score 2623.6; DB 1; Length 5314;  
Matches 3059; Conservative 0; Mismatches 579; Indels 28; Gaps 6;

QY 20 ATTTGGTTAAATTCGCGCTTAATTTTGTAAATCGTCATTTTAAACCAATAGGCC 79  
DB 1651 ATTTGGTTAAATTCGCGCTTAATTTTGTAAATCGTCATTTTAAACCAATAGGCC 79  
QY 80 GAAATCGGCAAAATTCCTTAATTAATCAAAAGAAATAGACGAGATAGGGTTAGTGT 1710  
DB 1711 GAAATCGGCAAAATTCCTTAATTAATCAAAAGAAATAGACGAGATAGGGTTAGTGT 139  
QY 140 CCAAGTTTGAACAAGATCCATTAATTAAGAAAGTGAATCTCCACGTCAAGAGGGGAAA 1770  
DB 1771 CCAAGTTTGAACAAGATCCATTAATTAAGAAAGTGAATCTCCACGTCAAGAGGGGAAA 199  
QY 200 ACCGCTTAATGAGGCGCATGCGCCACTACGTGAACCATCAACCTTAATCAAGTTTGGGG 1830  
DB 1831 ACCGCTTAATGAGGCGCATGCGCCACTACGTGAACCATCAACCTTAATCAAGTTTGGGG 259  
QY 260 TCGAGGTCGCGTAAGCACTAAATCGAAACCTTAAGAGAGCCCGCATTTAGAGCTTGA 1890  
DB 1891 TCGAGGTCGCGTAAGCACTAAATCGAAACCTTAAGAGAGCCCGCATTTAGAGCTTGA 1950  
QY 320 CCGGGAAAGCCGCGCAAGTGGCCAGAAAGGAAAGGAAAGGAAAGGAGCGGCGCT 379  
DB 1951 CCGGGAAAGCCGCGCAAGTGGCCAGAAAGGAAAGGAAAGGAAAGGAGCGGCGCT 2010  
QY 380 AGGGGCGTGGCAAGTGAAGCGGTCAAGCTGCGGTAACCAACACCCGCGCTTAAT 439  
DB 2011 AGGGGCGTGGCAAGTGAAGCGGTCAAGCTGCGGTAACCAACACCCGCGCTTAAT 2070  
QY 440 GCGCGCTACAGGCGCGGTCCCAATGCGCACTTCAAGCTGCGCACTGTTGGGAAAGGCGA 499  
DB 2071 GCGCGCTACAGGCGCGGTCCCAATGCGCACTTCAAGCTGCGCACTGTTGGGAAAGGCGA 2130  
QY 500 TCGGTGAGGCGCTTTCGCTAATTCGCGCACTGCGGCAAGGAGGATGTCGCAAGCGA 559  
DB 2131 TCGGTGAGGCGCTTTCGCTAATTCGCGCACTGCGGCAAGGAGGATGTCGCAAGCGA 2190  
QY 560 TTAAAGTTGGATACCGCAAGGTTTCCAGTCAAGACCTTGTAAACGACGCGCAAGTGA 619  
DB 2191 TTAAAGTTGGATACCGCAAGGTTTCCAGTCAAGACCTTGTAAACGACGCGCAAGTGA 2250  
QY 620 CCGGCGCTCGTCAATTCAGTTTGAACCGGTGAGGACGCGGAGACCTGCGGTGAAT 679  
DB 2251 TTGTAAATGCACTCACTAAGGCGCAATTTGAGGCTCCACCGCGGTGCGGTCTTGT 2310  
QY 680 GTGTAAATGCACTCACTAAGGCGCAATTTGAGGCTCCACCGCGGTGCGGTCTTGT 2370  
DB 2371 GTGTAAATGCACTCACTAAGGCGCAATTTGAGGCTCCACCGCGGTGCGGTCTTGT 799  
QY 740 GATTAACCTTAAGAAATTAATCAATTTTGAAGTGAAGGCTTAAGAAATTAATCAATTT 799  
DB 2371 GATTAACCTTAAGAAATTAATCAATTTTGAAGTGAAGGCTTAAGAAATTAATCAATTT 2370  
QY 800 TGAAGCATGAGATCTGAATACGACTCACTAAGGCGCAATTTGGGTACCGGCGCTT 859  
DB 2431 TGAAGCATGAGATCTGAATACGACTCACTAAGGCGCAATTTGGGTACCGGCGCTT 2430  
QY 860 GAGAGTCGACGATACG-----TAACTTGAATTCGAATTCCTGACGCGGGGAAATC 913  
DB 2491 GAGAGTCGACGATACG-----TAACTTGAATTCGAATTCCTGACGCGGGGAAATC 913  
QY 914 CACTAATTTAGAGGCGCGCAACCGCGGTGAGCTCAGCTTTTGTTCCTTAAGTGAAG 973  
DB 2551 CACTAATTTAGAGGCGCGCAACCGCGGTGAGCTCAGCTTTTGTTCCTTAAGTGAAG 2550  
QY 974 GGTAAATTAATCCAT-----GCGTCAATTTTACGACGACTAATCTTTTACGAGTTAA 1026  
DB 974 GGTAAATTAATCCAT-----GCGTCAATTTTACGACGACTAATCTTTTACGAGTTAA 1026

DB 2611 GATTCACCGCTTCAGATGACGACAGAACCAATGATTTGTAAATTCGAGAGAAATGCG 2670  
QY 1027 TCTAGCTGATGATGATGATTAATTCGCGGCTTTTTCGCGCTCAATTCGCCCAAGCT 1086  
DB 2671 CAGAGAAATCAATCAATTAAGATTTGAACCTTCTTGTGTCTTCAAGATGATTCCT 2730  
QY 1087 GCGCTATCTGGGATCGGGGAGAAAGAACCCGCTTTCCTCCGAGTTGAAGCGG 1146  
DB 2731 CATGTTTAAATGCGCTTGAAGATCCGCAAGCTTGAAGCTCAGGTGACCTTGA 2790  
QY 1147 CATGAAAGATTTGCCAGAGATGACTGCTGCTGCAATTAAGGAGGAAAGGACGCT 1206  
DB 2791 GATTCGCCAGAGCTCAGAGCTTAATTAATTTTCCGTAATTCAGGCTTCATGAT 2850  
QY 1207 TTACATGATGATTTCCGAGAGTGTGCGCATGACAGCTTAAAGGTGAATGTTCTTTC 1266  
DB 2851 GAGACAGGCGCTT---GAATGTTGACGGGATGAATTAATTAAGCAAGCGCAATTA 2908  
QY 1267 AGGCACTTGGGATACAGTTCGTCGCGCTTTTCGGAACAGTTCGGAATGTTGAGC 1326  
DB 2909 AACTCAACAG---AGCAGAAAGGAGGATATCCCAAAAGTCAAGCTTACATTA 2965  
QY 1327 CGAAGGCTACGCAACCTGAAACAAATACCGGCAAG-----CGGAATGCGGTGCGG 1380  
DB 2966 CAAAGCTCAAGCGACAGACGAGACAGAGGCGGTCAAGTACCAATCTTGTACTC 3025  
QY 1381 GTTGCAGATTAATGACAGCGGTGCGGCTGCGGATTAATTCGTACGAGAGAGCGGTAT- 1439  
DB 3026 GTTGCAGATTAATGACAGCGGTGCGGCTGCGGATTAATTCGTACGAGAGAGCGGTAT- 1439  
QY 1440 ---CTGCGTGGAGCGCGCAAAATGATGATATCCCGTGAATTCGCGGCGGCGC 1496  
DB 3086 CGACCTGAGGCGGCGCGGTACCGAGCTTGTGTTCCCTTAAAGAGGTTAATTCGA 3145  
QY 1497 GCTTGGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1556  
DB 3146 GCTTGGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3205  
QY 1557 CACACAAATTAAGAGCGGAGCAATTAAGTGAAGCTGCGGCTTAATGATGATGATGAT 1616  
DB 3206 CACACAAATTAAGAGCGGAGCAATTAAGTGAAGCTGCGGCTTAATGATGATGATGAT 3265  
QY 1617 AACTCAATTAATTTGCGTCTGCTCACTGCGCGCTTTCAGTGGGAAACCTGCTGTC 1676  
DB 3266 AACTCAATTAATTTGCGTCTGCTCACTGCGCGCTTTCAGTGGGAAACCTGCTGTC 3325  
QY 1677 AACTCAATTAATTTGCGTCTGCTCACTGCGCGCTTTCAGTGGGAAACCTGCTGTC 1736  
DB 3326 AACTCAATTAATTTGCGTCTGCTCACTGCGCGCTTTCAGTGGGAAACCTGCTGTC 3385  
QY 1737 CCGCTTCTGCGTCACTGACTGCTGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3445  
DB 3386 CCGCTTCTGCGTCACTGACTGCTGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3445  
QY 1797 CTAATTAAGGCGGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1856  
DB 3446 CTAATTAAGGCGGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3505  
QY 1857 TGTAGCAAAAGGCGCAAAAGGCGCAAGAACCGTAAAGGCGCGGTGCTGCTGCTT 1916  
DB 3506 TGTAGCAAAAGGCGCAAAAGGCGCAAGAACCGTAAAGGCGCGGTGCTGCTGCTT 1976  
QY 1917 TCGATAGGCTCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3625  
DB 3566 TCGATAGGCTCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3625  
QY 1977 GAAACCGGACAGAGCTTAATTAAGATCAAGGCGTTTCCCTGGAAGCTTCCCTGCGCT 2036  
DB 3626 GAAACCGGACAGAGCTTAATTAAGATCAAGGCGTTTCCCTGGAAGCTTCCCTGCGCT 3685  
QY 2037 GTTCTGTTCCGACCTGCGCTTACCGGATACCTGTCGCTTTCCTTCCGGAAGCG 2096  
DB 3686 GTTCTGTTCCGACCTGCGCTTACCGGATACCTGTCGCTTTCCTTCCGGAAGCG 3745

2097 TGGCGCTTTCATAGCTCAGCGCTAGGTATCTCAGTTGGTGTAGGTCGTTCCCTCA 2156  
 3146 TGGCGCTTTCATAGCTCAGCGCTAGGTATCTCAGTTGGTGTAGGTCGTTCCCTCA 3805  
 2157 AGCTGGGCTGTGTGACGACACCCCGTTACGCCGACCGCGGCGCTTATCCGGTAAT 2216  
 3806 AGCTGGGCTGTGTGACGACACCCCGTTACGCCGACCGCGGCGCTTATCCGGTAAT 3865  
 2217 ATCGCTTGTAGTCCACCCCGTTAGACAGACTTATCCGCACTGACAGACCACTGTGTA 2276  
 3866 ATCGCTTGTAGTCCACCCCGTTAGACAGACTTATCCGCACTGACAGACCACTGTGTA 3925  
 2277 ACAGATTAGCAGACGAGAGTATGTAGGCGGTGTACAGAGTTCTTGAAGTGTGGCTTA 2336  
 3926 ACAGATTAGCAGACGAGAGTATGTAGGCGGTGTACAGAGTTCTTGAAGTGTGGCTTA 3985  
 2337 ACTACGGCTACACTGAGAGACAGATTTGTATCTGCGCTGTGCTGAAGCCAGTTACT 2396  
 3986 ACTACGGCTACACTGAGAGACAGATTTGTATCTGCGCTGTGCTGAAGCCAGTTACT 4045  
 2397 TCGGAAAAAGTTGTAGCTCTTGTATCCGCAACAAACACCGCTGGTAGCGGTGTT 2456  
 4046 TCGGAAAAAGTTGTAGCTCTTGTATCCGCAACAAACACCGCTGGTAGCGGTGTT 4105  
 2457 TTTTGTGTGCAAGCAGCAGATTACGCGCAAAAAAGATCTCAAGAGATCTTTGA 2516  
 4106 TTTTGTGTGCAAGCAGCAGATTACGCGCAAAAAAGATCTCAAGAGATCTTTGA 4165  
 2517 TCTTTTCAAGGGGTCTGACGCTCAGTGAAGGAAACGAAACCTCAGTGAAGGATTTGTGCA 2576  
 4166 TCTTTTCAAGGGGTCTGACGCTCAGTGAAGGAAACCTCAGTGAAGGATTTGTGCA 4225  
 2577 TGAATTATCAAAAAGATCTTCACTAGATCTTTTAAATTTAAATGAAGTTTAAAT 2636  
 4226 TGAATTATCAAAAAGATCTTCACTAGATCTTTTAAATTTAAATGAAGTTTAAAT 4285  
 2637 CAATCTAAAGTATATATAGATTAATTTGTCTGACAGATTCCAAATGCTTATCACTGAGG 2696  
 4286 CAATCTAAAGTATATATAGATTAATTTGTCTGACAGATTCCAAATGCTTATCACTGAGG 4345  
 2697 CACCTATCTACGCGATCTGTCTATTTGCTTCAATGATGCTGATCCCGCGCGGT 2756  
 4346 CACCTATCTACGCGATCTGTCTATTTGCTTCAATGATGCTGATCCCGCGCGGT 4405  
 2757 AGATACTACAGATACGGAGGCGCTTACCATCTGCGCCCAAGTCTGCAATGATACCGCGAG 2816  
 4406 AGATACTACAGATACGGAGGCGCTTACCATCTGCGCCCAAGTCTGCAATGATACCGCGAG 4465  
 2817 ACCACGCTCACCGGCTCCAGATTTATCAGCAATTAACAGCCAGCCGGAAGGCGGAGC 2876  
 4466 ACCACGCTCACCGGCTCCAGATTTATCAGCAATTAACAGCCAGCCGGAAGGCGGAGC 4525  
 2877 GCAGAGGTGCTCTGCACTTTATCCGCTCCATCAGTCTTAAATTTGTCGCGGAGAG 2936  
 4526 GCAGAGGTGCTCTGCACTTTATCCGCTCCATCAGTCTTAAATTTGTCGCGGAGAG 4585  
 2937 CTAGAGTAAAGTATGCTGCGCAGTAAATTTGCGCAAGTGTGTCATTTGCTACAGGCA 2996  
 4586 CTAGAGTAAAGTATGCTGCGCAGTAAATTTGCGCAAGTGTGTCATTTGCTACAGGCA 4645  
 2997 TCGTGTGTGCAAGCTGCTGCTTTGTGTATGCTTCACTGCTCGGTTCCCAAGATCAA 3056  
 4646 TCGTGTGTGCAAGCTGCTGCTTTGTGTATGCTTCACTGCTCGGTTCCCAAGATCAA 4705  
 3057 GCGAGTTACATGATCCCGCATGTTGTGCAAAAAAGGGTTAGCTCTTGGTCTCCGA 3116  
 4706 GCGAGTTACATGATCCCGCATGTTGTGCAAAAAAGGGTTAGCTCTTGGTCTCCGA 4765  
 3117 TCGTGTGCAAGTAAAGTTGGCCGAGTGTATCACTAGTATTTGACAGACATGCAATA 3176  
 4766 TCGTGTGCAAGTAAAGTTGGCCGAGTGTATCACTAGTATTTGACAGACATGCAATA 4825

3177 ATTCCTTACTGTATGCAATCCGTAAGATGCTTTCTGTAGACGTGTAGTATCAACCA 3236  
 4826 ATTCCTTACTGTATGCAATCCGTAAGATGCTTTCTGTAGACGTGTAGTATCAACCA 4885  
 3237 AGTCAATCTGAGAAATAGTATGAGCGGACGAGTGTGCTTGTGCGCGCTCAATACGAG 3296  
 4886 AGTCAATCTGAGAAATAGTATGAGCGGACGAGTGTGCTTGTGCGCGCTCAATACGAG 4945  
 3297 ATATATCCGCGCCATAGCAGAACTTTAAAGTCTCATCATTTGAAACGTTCTTCCG 3356  
 4946 ATATATCCGCGCCATAGCAGAACTTTAAAGTCTCATCATTTGAAACGTTCTTCCG 5005  
 3357 GCGCAAACTCTCAAGATCTTACCGCTGTGAGATCCAGTTGATGTAACCATTCGTG 3416  
 5006 GCGCAAACTCTCAAGATCTTACCGCTGTGAGATCCAGTTGATGTAACCATTCGTG 5065  
 3417 CACCAACTGATCTTACAGCATCTTTTACTTACCAAGGCTTCTGCGTGAAGCAAAACG 3476  
 5066 CACCAACTGATCTTACAGCATCTTTTACTTACCAAGGCTTCTGCGTGAAGCAAAACG 5125  
 3477 GAAGCAAAATGCGCAAAAGGAAATAGCGCGACACGAAATGTTGAATCTCATAC 3536  
 5126 GAAGCAAAATGCGCAAAAGGAAATAGCGCGACACGAAATGTTGAATCTCATAC 5185  
 3537 TCTTCCTTTTCAATATATTAAGCATTTATCAGGTTATTTGCTCATAGCGGATACA 3596  
 5186 TCTTCCTTTTCAATATATTAAGCATTTATCAGGTTATTTGCTCATAGCGGATACA 5245  
 3597 TATTGAATGATTTAGAAATTAACCAATAGGGGTTCCGCGCATTTCCCGCAAAAG 3656  
 5246 TATTGAATGATTTAGAAATTAACCAATAGGGGTTCCGCGCATTTCCCGCAAAAG 5305  
 3657 TGCCAC 3662.  
 5306 TGCCAC 5311

RESULT 3  
 US-08-646-538-6  
 / Sequence 6, Application US/08646538  
 / Patent No. 6027881  
 GENERAL INFORMATION:  
 APPLICANT: Pavlakis, George N.  
 APPLICANT: Gaetanaris, George A.  
 APPLICANT: Stauber, Roland H.  
 APPLICANT: Vournakis, John N.  
 TITLE OF INVENTION: Mutant Aequorea victoria Fluorescent  
 TITLE OF INVENTION: Proteins Having Increased Cellular Fluorescence  
 NUMBER OF SEQUENCES: 37  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Townsend and Townsend and Crew LLP  
 STREET: Two Embarcadero Center, 8th Floor  
 CITY: San Francisco  
 STATE: California  
 COUNTRY: USA  
 ZIP: 94111-3834  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/646,538  
 FILING DATE: No. 6027881 yet assigned  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Weber, Kenneth A.  
 REGISTRATION NUMBER: 31,677  
 REFERENCE/DOCKET NUMBER: 015280-249000  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (415) 576-0200  
 TELEFAX: (415) 576-0300  
 INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:  
LENGTH: 3699 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA  
FEATURE:  
NAME/KEY:  
LOCATION: 1..3699  
OTHER INFORMATION: /note= "pBSGPR"

US-08-646-538-6

Query Match 71.6%; Score 2620.6; DB 3; Length 3699;  
Best Local Similarity 83.3%; Pred. No. 0;  
Matches 3078; Conservative 0; Mismatches 579; Indels 40; Gaps 7;

QY 3 AATTGTAGCGTTAATTTTGTAAATTCGCGTTAAATTTGTAAATTCAGCTCAT 62  
DB 3 AATTGTAAACGTTAATTTTGTAAATTCGCGTTAAATTTGTAAATTCAGCTCAT 62  
QY 63 TTTTAAACCAATAGCCGAAATCGCAAAATCCCTTAATTAATCAAAAGAAATGACGAGA 122  
DB 63 TTTTAAACCAATAGCCGAAATCGCAAAATCCCTTAATTAATCAAAAGAAATGACGAGA 122  
QY 123 TAGGTTGAGTTGTTTCCAGTTTGGACAAGATCCCTTAATTAATCAAAAGAAATGACGAGA 122  
DB 123 TAGGTTGAGTTGTTTCCAGTTTGGACAAGATCCCTTAATTAATCAAAAGAAATGACGAGA 122  
QY 183 ACCTCAAAAGGGGCAAAAACCGCTCTATCGAGCGATGCGCCACTACGTGAACATCACTCA 182  
DB 183 ACCTCAAAAGGGGCAAAAACCGCTCTATCGAGCGATGCGCCACTACGTGAACATCACTCA 182  
QY 243 AATCAAGTTTTTGGGGTCAGAGTGCCTTAAGCACTAAATCGGAAACCTTAAGGAGACC 242  
DB 243 AATCAAGTTTTTGGGGTCAGAGTGCCTTAAGCACTAAATCGGAAACCTTAAGGAGACC 242  
QY 303 CCCGATTTAGACTTGAACGAGGAAACCGCGCAACCTGCGCAAGAAAGAGGAAAGAAAG 302  
DB 303 CCCGATTTAGACTTGAACGAGGAAACCGCGCAACCTGCGCAAGAAAGAGGAAAGAAAG 302  
QY 363 CGAAGAGAGCGGGGCGTAAAGGCGCTGCGCAAGTGAACGCGTCAACGCTGGCGTAAACCA 362  
DB 363 CGAAGAGAGCGGGGCGTAAAGGCGCTGCGCAAGTGAACGCGTCAACGCTGGCGTAAACCA 362  
QY 423 CACCCGCGCGCTTAATGCGCGCTACAGGCGCGTCAACGCTGGCGTAAACCA 422  
DB 423 CACCCGCGCGCTTAATGCGCGCTACAGGCGCGTCAACGCTGGCGTAAACCA 422  
QY 480 GCAACTGTTGGGAAGGCGGATCGGTGCGGCTCTTCTGCTAATTAACCCAGCTGGCGAAG 479  
DB 480 GCAACTGTTGGGAAGGCGGATCGGTGCGGCTCTTCTGCTAATTAACCCAGCTGGCGAAG 479  
QY 540 GGGGATGTCGCAAGGCGGATCGGTGCGGCTCTTCTGCTAATTAACCCAGCTGGCGAAG 539  
DB 540 GGGGATGTCGCAAGGCGGATCGGTGCGGCTCTTCTGCTAATTAACCCAGCTGGCGAAG 539  
QY 599 GGGGATGTCGCAAGGCGGATTAAGTTAGTTAGTTAGTTAGTTAGTTAGTTAGTTAGTTAG 599  
DB 599 GGGGATGTCGCAAGGCGGATTAAGTTAGTTAGTTAGTTAGTTAGTTAGTTAGTTAGTTAG 599  
QY 602 GTAAACGACGCGCAGATGATGTAAT--ACACTCACTATAGGCGCAATGGGTACCGG 602  
DB 602 GTAAACGACGCGCAGATGATGTAAT--ACACTCACTATAGGCGCAATGGGTACCGG 602  
QY 659 GGCAGACTCGCGGGAATGTTTTCACAGCGTATGAGAGAGTGAAGTGTGCA 659  
DB 659 GGCAGACTCGCGGGAATGTTTTCACAGCGTATGAGAGAGTGAAGTGTGCA 659  
QY 719 GCGTGC---GAACACGACGACTAGATTAAACCTTAAAGATTAATCATATTTGACGTAAG 719  
DB 719 GCGTGC---GAACACGACGACTAGATTAAACCTTAAAGATTAATCATATTTGACGTAAG 719  
QY 776 CATGCCATGTGTAATCCGACGCTGTAACAACTCAAGAGAGACATGCGTCTCTT 780  
DB 776 CATGCCATGTGTAATCCGACGCTGTAACAACTCAAGAGAGACATGCGTCTCTCTT 780  
QY 836 TTAAGATATCATGCGTAAATTAAGCATGAGATGTAATTAAGCATGCGTCTCTT 836  
DB 836 TTAAGATATCATGCGTAAATTAAGCATGAGATGTAATTAAGCATGCGTCTCTCTT 836  
QY 839 GAATTGGGTACCGGGCCCCCTCGAGGTGCAAGGTATGATTAAGTTGATTCGATTC 896  
DB 839 GAATTGGGTACCGGGCCCCCTCGAGGTGCAAGGTATGATTAAGTTGATTCGATTC 896

DB 840 GAAAGAGGCGCATGCGCAATGGAATTTTGTATATGCTGCTAGTTGAACGCTTC 899  
QY 897 CTCACACCGCGGGGATTCATAGTT---CTAGAGCGCGCCACCGCGGTGAGCTCCA 952  
DB 900 CATCTTCAATGTTGTGTCTAAATTTTGAAGTTAACTTTGATTCATCTTTTGTGTCTG 959  
QY 953 GCTTTTGTCCCTTAGTGAAGGTTAATTAATGATCCATGCGTCAATTTTAAAGCATAT 1012  
DB 960 CCATATGTATATCATGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1019  
QY 1013 CTTTCTAGAGTTATCTAGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1072  
DB 1020 CATCTTCTTAAATTAATTAATCTTTTAACTGATGATGATGATGATGATGATGATGAT 1079  
QY 1073 TCATCGCCCAAGCTGCGCTATCTGAGGATGCGGAGAGAGAGAGAGAGAGAGAGAGAG 1139  
DB 1080 ACTTGAATTCAGACGAGT 1139  
QY 1133 CGAGTTGAAGCGGCGATGGAAGAGTTTCCGAGAGATGATGCTGCTGATGATGATGATGAT 1192  
DB 1140 GTACATTAACCTTCGCGCATGCGATGCTGTAAGAAAGATGATGATGATGATGATGATGAT 1199  
QY 1193 GCGAAACCGCATGTTTACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1252  
DB 1200 ATCTTGAAGAGATTTGAACATTAAGAAAGATGATGATGATGATGATGATGATGATGAT 1259  
QY 1253 TGAATGTTGTTGAGGCGACCTGGGATACCACTGCTGCGGCTTTTCCGACAGACTT 1312  
DB 1260 GTAGTTTTCAGTGTGCAATTAATTAAGGTTAAGTTTCCGATGATGATGATGATGAT 1319  
QY 1313 CGGATGCTGAGCGCGGAGCGCATGCAACCGCAACCAATACCGGCGCACCGGAACTG 1372  
DB 1320 CACCTTCTCACTGACGAAATTTGTGCGCATTAATCACTATTAATCAACCAAGAA 1379  
QY 1373 CCGTGC-----GATGCGAGATTAATGACAGCGTGGCGCTGGAGATAT 1418  
DB 1380 TTGGGACACTCCAGTGAAGAGTTCTTCTTCTGCTAGCCATTTTGGCGCATGCAAT 1439  
QY 1419 TACGTCGAGAGAGCGGATATCTGCTGATGCGGAGAAATGAGATGATGATGATGATGAT 1478  
DB 1440 TCCGACGCGCGGAGATCACTAATGTTGAGGCGCGGACCGCGGTGAGCTCAAC 1499  
QY 1479 GAGTTACCGCGGCG-----CGGCTGGCGGATTAATGATGATGATGATGATGATGAT 1525  
DB 1500 TTTTGTCCCTTATGAGGCTTAATTCAGAGCTTGAATGATGATGATGATGATGATGAT 1559  
QY 1526 CCGTGTGAATTTGTTATCGCTCAGATTTCACACATATGAGAGCGGAAAG 1585  
DB 1560 CCGTGTGAATTTGTTATCGCTCAGATTTCACACATATGAGAGCGGAAAG 1619  
QY 1586 TGTAAAGCTGGGGTGTCTATAGTGAATTAATTAATGAGTTTGGCTCACTG 1645  
DB 1620 TGTAAAGCTGGGGTGTCTATAGTGAATTAATTAATGAGTTTGGCTCACTG 1679  
QY 1646 CCGCTTTTCAAGTGGGAAACCTGTGCGGACGCTGCTTAATTAATGAGTTTGGCTCACTG 1705  
DB 1680 CCGCTTTTCAAGTGGGAAACCTGTGCGGACGCTGCTTAATTAATGAGTTTGGCTCACTG 1739  
QY 1706 GGGAGAGCGGTTTGGGATTTGGGCTCTTCCGCTTCTGCTCACTGAGCTGCGC 1765  
DB 1740 GGGAGAGCGGTTTGGGATTTGGGCTCTTCCGCTTCTGCTCACTGAGCTGCGC 1799  
QY 1766 TCGGTGTTGCGCTGCGGCGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1825  
DB 1800 TCGGTGTTGCGCTGCGGCGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1859  
QY 1826 ACAGATCAGGGGATTAAGAGAGAAAGATGATGATGATGATGATGATGATGATGATGATGAT 1885  
DB 1860 ACAGATCAGGGGATTAAGAGAGAAAGATGATGATGATGATGATGATGATGATGATGATGAT 1919  
QY 1886 AACGTAAGAGCGCGGTTGCTGCTGTTTTCATAGGCTCCGCGCTGACGAGCAT 1945

Db 1920 AACGTAAGAGCGCGGTGCTGGCGTTTTCATAGAGTCCGCCCTGAGAGCAT 1979  
Qy 1946 CACAAAATCGACGCTCAAGTCAGAGGTGGGAAACCCGACAGCATTAAGATACAG 2005  
Db 1980 CACAAAATCGACGCTCAAGTCAGAGGTGGGAAACCCGACAGCATTAAGATACAG 2039  
Qy 2006 GCGTTTCCCGTGGAGGTCCCTGTGGGCTCTCCGTGTCGAGCCCTGCGCTTAACCGGA 2065  
Db 2040 GCGTTTCCCGTGGAGGTCCCTGTGGGCTCTCCGTGTCGAGCCCTGCGCTTAACCGGA 2099  
Qy 2066 TACGTGTCGCGCTTCTCTCTCGGGAAGCGTGGCGCTTCTCATAGTCAAGCTGAGG 2125  
Db 2100 TACGTGTCGCGCTTCTCTCTCGGGAAGCGTGGCGCTTCTCATAGTCAAGCTGAGG 2159  
Qy 2126 TATCTCAAGTTGGGTAGGTGTTGCTGCTCAAGCTGGGCTGTGTCAGAAACCCCGCTT 2185  
Db 2160 TATCTCAAGTTGGGTAGGTGTTGCTGCTCAAGCTGGGCTGTGTCAGAAACCCCGCTT 2219  
Qy 2186 CAGCCCGACGCTGGGCTTATCCGGGTAACTATGCTTGAAGTCCAGCCCGTAAAGACAC 2245  
Db 2220 CAGCCCGACGCTGGGCTTATCCGGGTAACTATGCTTGAAGTCCAGCCCGTAAAGACAC 2279  
Qy 2246 GACTTATCGCCATCGGACAGCACTGTAAACAGATTAGCAGAGCGAGGTATGTAGGC 2305  
Db 2280 GACTTATCGCCATCGGACAGCACTGTAAACAGATTAGCAGAGCGAGGTATGTAGGC 2339  
Qy 2306 GGTGCTACAGAGTTCTTGAAGTGGTGGCTTAACTACGGCTACACTAGAGACAGTATTT 2365  
Db 2340 GGTGCTACAGAGTTCTTGAAGTGGTGGCTTAACTACGGCTACACTAGAGACAGTATTT 2399  
Qy 2366 GGTATCGCGCTGCTGTAAGCACTTACCTCGGAAAAGAGTTGAGTCTTTGATCC 2425  
Db 2400 GGTATCGCGCTGCTGTAAGCACTTACCTCGGAAAAGAGTTGAGTCTTTGATCC 2459  
Qy 2426 GGGCAACAAACACCGCTGTGAGCGGTGTTTTTTTGTTCAGACAGACATTAACGGC 2485  
Db 2460 GGGCAACAAACACCGCTGTGAGCGGTGTTTTTTTGTTCAGACAGACATTAACGGC 2519  
Qy 2486 AGAAAAAAGAGATCTCAAGAAATCTTTGATCTTTTCTACGGGGCTGACGCTCAGTGG 2545  
Db 2520 AGAAAAAAGAGATCTCAAGAAATCTTTGATCTTTTCTACGGGGCTGACGCTCAGTGG 2579  
Qy 2546 AAGAAAACTCAGCTTAAAGGATTTGGTCAATGAGATTATCAAAAAGGATCTTACCCAG 2605  
Db 2580 AAGAAAACTCAGCTTAAAGGATTTGGTCAATGAGATTATCAAAAAGGATCTTACCCAG 2639  
Qy 2606 ATCTTTTAAATTAAGATTAAGATTTTAAATCAATCTAAAGTATATGATTAACCTGG 2665  
Db 2640 ATCTTTTAAATTAAGATTAAGATTTTAAATCAATCTAAAGTATATGATTAACCTGG 2699  
Qy 2666 TCTGACAGTTACCAATGCTTAATCAGTGAAGGACCTATCTGACGAGTCTGTATTTCTG 2725  
Db 2700 TCTGACAGTTACCAATGCTTAATCAGTGAAGGACCTATCTGACGAGTCTGTATTTCTG 2759  
Qy 2726 TCAATCCATAGTTGGCTGATCTCCCGTGTGTAAGTAATAAGTAAAGGAGGCTTACCA 2785  
Db 2760 TCAATCCATAGTTGGCTGATCTCCCGTGTGTAAGTAATAAGTAAAGGAGGCTTACCA 2819  
Qy 2786 TCTGGCCCAAGTGTGCAATGATACCGGAGAACCCACGCTCAGCGGCTCCAGATTATCA 2845  
Db 2820 TCTGGCCCAAGTGTGCAATGATACCGGAGAACCCACGCTCAGCGGCTCCAGATTATCA 2879  
Qy 2846 GCAATTAACCAAGCCAGCGGAAAGGCGGAGCGCAAGAGTGTCTGCAACTTATCCGCC 2905  
Db 2880 GCAATTAACCAAGCCAGCGGAAAGGCGGAGCGCAAGAGTGTCTGCAACTTATCCGCC 2939  
Qy 2906 TCCATCCAGTCTATTAATGTTGCGGGAAGCTAGAGTAATAGTCCCGCATTAATAGT 2965  
Db 2940 TCCATCCAGTCTATTAATGTTGCGGGAAGCTAGAGTAATAGTCCCGCATTAATAGT 2999  
Qy 2966 TTGGGCAACGTTGTTGGCATTTGCTACAGGCACTGAGTGTACCGCTCGCTTTGGTATG 3025  
Db 3000 TTGGGCAACGTTGTTGGCATTTGCTACAGGCACTGAGTGTACCGCTCGCTTTGGTATG 3059

Qy 3026 GCTTCATTACACTCCGTTCCCAACGATCAAGGCGAGTATACATGATCCCCCATGTTGTC 3085  
Db 3060 GCTTCATTACACTCCGTTCCCAACGATCAAGGCGAGTATACATGATCCCCCATGTTGTC 3119  
Qy 3086 AAAAAAGCGGTAGCTCCCTCGGTCTCCGATCCGTTGTCAGAAAGTAAAGTGGCCGACGTG 3145  
Db 3120 AAAAAAGCGGTAGCTCCCTCGGTCTCCGATCCGTTGTCAGAAAGTAAAGTGGCCGACGTG 3179  
Qy 3146 TTATCACTCATGTTATAGCAGCACTGATTAATCTCTTACTGTATGCCATCCGTAGA 3205  
Db 3180 TTATCACTCATGTTATAGCAGCACTGATTAATCTCTTACTGTATGCCATCCGTAGA 3239  
Qy 3206 TGCCTTTCTGATCTGTTGAGTACTCAACCAAGTCTTGAAGTATGTATGGCGCA 3265  
Db 3240 TGCCTTTCTGATCTGTTGAGTACTCAACCAAGTCTTGAAGTATGTATGGCGCA 3299  
Qy 3286 CCGAGTGTCTTTCGCGCGGCTCAATACGGGATTAATACCGGCAATAGCAGAACTTTA 3325  
Db 3300 CCGAGTGTCTTTCGCGCGGCTCAATACGGGATTAATACCGGCAATAGCAGAACTTTA 3359  
Qy 3326 AAAGTGTCACTATGGAAGAAAGTTCTTCGCGGCGGAAACTCTCAAGATCTTACCGCTG 3385  
Db 3360 AAAGTGTCACTATGGAAGAAAGTTCTTCGCGGCGGAAACTCTCAAGATCTTACCGCTG 3419  
Qy 3386 TTGAGATCCAGTGTGATGTAACCACTGCTGCAACCACTGATCTTACGATCTTTTACT 3445  
Db 3420 TTGAGATCCAGTGTGATGTAACCACTGCTGCAACCACTGATCTTACGATCTTTTACT 3479  
Qy 3446 TTCAACGACGTTTCGCGGTGAGCAAAAACAGGAAGGCAAAATGCGGCAAAAAGGGAATA 3505  
Db 3480 TTCAACGACGTTTCGCGGTGAGCAAAAACAGGAAGGCAAAATGCGGCAAAAAGGGAATA 3539  
Qy 3506 AGGCGCACAGCAAAATGTTGAATACATATCTCTTCTTTTCAATATTAATGAAGCAT 3565  
Db 3540 AGGCGCACAGCAAAATGTTGAATACATATCTCTTCTTTTCAATATTAATGAAGCAT 3599  
Qy 3566 TATCAGGTTATTTGTCATAGAGGAGTATCTTGAATATTTGAAGTATTAAGAAATTAACA 3625  
Db 3600 TATCAGGTTATTTGTCATAGAGGAGTATCTTGAATATTTGAAGTATTAAGAAATTAACA 3659  
Qy 3626 ATAGGGGTTCGCGCACATTTCCCGAAAGTGCAC 3662  
Db 3660 ATAGGGGTTCGCGCACATTTCCCGAAAGTGCAC 3696

RESULT 4  
US-09-503-222-6  
Sequence 6, Application US/09503222  
Patent No. 6265548  
GENERAL INFORMATION:  
APPLICANT: Pavlakis, George N.  
APPLICANT: Galanaris, George A.  
APPLICANT: Stauber, Roland H.  
APPLICANT: Vournakis, John N.  
TITLE OF INVENTION: Mutant Aegeorea victoria Fluorescent  
TITLE OF INVENTION: Proteins Having Increased Cellular Fluorescence  
NUMBER OF SEQUENCES: 37  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, 8th Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patencin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/503,222  
FILING DATE:





Db 1740 GGGAGAGCGGTTTTCGTAATTGAGCGCTCTTCCGCTCCTCAGTCACTGACTCGCTGCGC 1799  
Qy 1766 TCGGTCGTTCCGCTGCGGCGAGCGGCTATCAGTCTCACTAAAGCGGTAATAGGTTATCC 1825  
Db 1800 TCGGTCGTTCCGCTGCGGCGAGCGGCTATCAGTCTCACTAAAGCGGTAATAGGTTATCC 1859  
Qy 1826 ACAGATCAGGGGATATACGACAGAAAGAAATGATGAGCAAAAGGCGACAAAGGCGAG 1885  
Db 1860 ACAGATCAGGGGATATACGACAGAAAGAAATGATGAGCAAAAGGCGACAAAGGCGAG 1919  
Qy 1886 AACGTTAAAAAGCGCGCTGTGCGGTTTTTTCATAGGCTCCGCGCCCTGACAGCAT 1945  
Db 1920 AACGTTAAAAAGCGCGCTGTGCGGTTTTTTCATAGGCTCCGCGCCCTGACAGCAT 1979  
Qy 1946 CACAAAAATCGACGCTCAAGTCAAGGTGCGGAAACCCGACAGGACTATTAAGATACAG 2005  
Db 1980 CACAAAAATCGACGCTCAAGTCAAGGTGCGGAAACCCGACAGGACTATTAAGATACAG 2039  
Qy 2006 GCGTTTTCCCGTGGAAAGCTCCCTCGTGGCTCTCCTGTTCCGACCCCTGCGCTTACCGGA 2065  
Db 2040 GCGTTTTCCCGTGGAAAGCTCCCTCGTGGCTCTCCTGTTCCGACCCCTGCGCTTACCGGA 2099  
Qy 2066 TACCTGTCGCGCTTCTCTCCTTCCGGAAGCGTGGCGCTTCTCATAGCTCAGCTGTAGG 2125  
Db 2100 TACCTGTCGCGCTTCTCTCCTTCCGGAAGCGTGGCGCTTCTCATAGCTCAGCTGTAGG 2159  
Qy 2126 TATCTCAAGTTGGGTGATAGTGTGCTGCTCCAAAGCTGGGCTGTGTCAGAAACCCCGCTT 2185  
Db 2160 TATCTCAAGTTGGGTGATAGTGTGCTGCTCCAAAGCTGGGCTGTGTCAGAAACCCCGCTT 2219  
Qy 2186 CAGCCCGACGCGCTGCGCTTATCCGGTAATATGCTTGTGATCCAAACCGGTAAGACAC 2245  
Db 2220 CAGCCCGACGCGCTGCGCTTATCCGGTAATATGCTTGTGATCCAAACCGGTAAGACAC 2279  
Qy 2246 GACTTATCGCACTGCGACAGCACTGTAACAGATTAAGACAGGATATGTAGGC 2305  
Db 2280 GACTTATCGCACTGCGACAGCACTGTAACAGATTAAGACAGGATATGTAGGC 2339  
Qy 2306 GGTCTCAAGATTTCTGAAGTGTGCTTAATACGCTCACTAGAAAGACAGTATTT 2365  
Db 2340 GGTCTCAAGATTTCTGAAGTGTGCTTAATACGCTCACTAGAAAGACAGTATTT 2399  
Qy 2366 GGTATCGCGCTGCTGTAAGCACTTCCGAAAAAGATTTGGATGCTTGAATCC 2425  
Db 2400 GGTATCGCGCTGCTGTAAGCACTTCCGAAAAAGATTTGGATGCTTGAATCC 2459  
Qy 2426 GCGAAACAAACACGCTGTAAGCGTGTGTTTTTGTGTCAGACAGCAATTAACGCGC 2485  
Db 2460 GCGAAACAAACACGCTGTAAGCGTGTGTTTTTGTGTCAGACAGCAATTAACGCGC 2519  
Qy 2486 AGAAAAAAGGATCTCAAGAAATCTTTGATCTTTTCTACGCGGCTGACGCTCAAGTG 2545  
Db 2520 AGAAAAAAGGATCTCAAGAAATCTTTGATCTTTTCTACGCGGCTGACGCTCAAGTG 2579  
Qy 2546 AACGAAAACTCAGCTTAAGGATTTGGTCAATGATTAATCAAAAAGGATCTTCAACCTAG 2605  
Db 2580 AACGAAAACTCAGCTTAAGGATTTGGTCAATGATTAATCAAAAAGGATCTTCAACCTAG 2639  
Qy 2606 ATCTTTTAAATTAATGAAGTTTAAATCAATCTAAAGTATATGATTAACCTTGG 2665  
Db 2640 ATCTTTTAAATTAATGAAGTTTAAATCAATCTAAAGTATATGATTAACCTTGG 2699  
Qy 2666 TCTGACAGTTACCAATGCTTAATCAGTAGGCACTTCTCAGCGATCTGTCTATTTCTGT 2725  
Db 2700 TCTGACAGTTACCAATGCTTAATCAGTAGGCACTTCTCAGCGATCTGTCTATTTCTGT 2759  
Qy 2726 TCTATCAGTTAGTCTGACCTCCGCTGCTGATGATTAATCAATGATTAACGAGGCTTACCA 2785  
Db 2760 TCTATCAGTTAGTCTGACCTCCGCTGCTGATGATTAATCAATGATTAACGAGGCTTACCA 2819  
Qy 2786 TCTGAGCCCAAGTGTGCAATGATACCGGAGAACCAAGCTCAGCGGCTCCAGATTTATCA 2845  
Db 2820 TCTGAGCCCAAGTGTGCAATGATACCGGAGAACCAAGCTCAGCGGCTCCAGATTTATCA 2879

Qy 2846 GCAATTAACCAAGCCAGCCGGAAGGCGGACAGAAAGTGGTCTTGCACACTTTATCCGCC 2905  
Db 2880 GCAATTAACCAAGCCAGCCGGAAGGCGGACAGAAAGTGGTCTTGCACACTTTATCCGCC 2939  
Qy 2906 TCCATCCAGTCTATTAATTTGTGCGGGAAGCTAGAGTACATGATTCGCGAGTTAATGT 2965  
Db 2940 TCCATCCAGTCTATTAATTTGTGCGGGAAGCTAGAGTACATGATTCGCGAGTTAATGT 2999  
Qy 2966 TTGCGCAAGTGTGCGCATTTGCTACAGAGCATGCTGTGTCACGCTGCTGTGGTATG 3025  
Db 3000 TTGCGCAAGTGTGCGCATTTGCTACAGAGCATGCTGTGTCACGCTGCTGTGGTATG 3059  
Qy 3026 GCTTATTCAGCTCCGCTTCCCAACGATCAAGGGAAGTTAATGATCCCAATGTTGTC 3085  
Db 3060 GCTTATTCAGCTCCGCTTCCCAACGATCAAGGGAAGTTAATGATCCCAATGTTGTC 3119  
Qy 3086 AAAAAAGGCTTACCTCCTCGGTCTCCGATCGTTGTCAGAAAGTGGCGCAGTG 3145  
Db 3120 AAAAAAGGCTTACCTCCTCGGTCTCCGATCGTTGTCAGAAAGTGGCGCAGTG 3179  
Qy 3146 TTATCACTCATGTTATGACAGCACTGCAATATTTCTTACTGTCAATGCCATCCGTAGA 3205  
Db 3180 TTATCACTCATGTTATGACAGCACTGCAATATTTCTTACTGTCAATGCCATCCGTAGA 3239  
Qy 3206 TGCTTTTCTGTGATCGTGTGATCTCAACCAAGTCAATTTGCAAAATGATGAGCGCA 3265  
Db 3240 TGCTTTTCTGTGATCGTGTGATCTCAACCAAGTCAATTTGCAAAATGATGAGCGCA 3299  
Qy 3266 CCGAGTGTCTTCCGCGCGCTCAATACGGGATTAATCCGCGCAATAGAGAACTTTA 3325  
Db 3300 CCGAGTGTCTTCCGCGCGCTCAATACGGGATTAATCCGCGCAATAGAGAACTTTA 3359  
Qy 3326 AAAGTGTCTCATTTGAAAAAGCTTCTTCCGCGGGAAGAACTTCAAGATCTTACCGCTG 3385  
Db 3360 AAAGTGTCTCATTTGAAAAAGCTTCTTCCGCGGGAAGAACTTCAAGATCTTACCGCTG 3419  
Qy 3386 TTGAGATTCAGATTTGATGATTAACCACTGCTGACCCAACTGATCTTCAAGATCTTTTAACT 3445  
Db 3420 TTGAGATTCAGATTTGATGATTAACCACTGCTGACCCAACTGATCTTCAAGATCTTTTAACT 3479  
Qy 3446 TTCAACGACGCTTCTGCGTGAAGAAAAAGGAAAGGCAAAATGCGGAAAAAGGAGATA 3505  
Db 3480 TTCAACGACGCTTCTGCGTGAAGAAAAAGGAAAGGCAAAATGCGGAAAAAGGAGATA 3539  
Qy 3506 AGGCGACACGGAATGTTGAATACATCTCTTCTTTTCAATATTAATTTGAAGATTT 3565  
Db 3540 AGGCGACACGGAATGTTGAATACATCTCTTCTTTTCAATATTAATTTGAAGATTT 3599  
Qy 3566 TATCAGGCTTAATGCTCATAGAGCGGATACATATTTGAATGATTTAGAAAAATTAACAA 3625  
Db 3600 TATCAGGCTTAATGCTCATAGAGCGGATACATATTTGAATGATTTAGAAAAATTAACAA 3659  
Qy 3626 ATAGGGGTTCCGCGCACTTCCCGAAAAAGTGCAC 3662  
Db 3660 ATAGGGGTTCCGCGCACTTCCCGAAAAAGTGCAC 3696

RESULT 5  
US-09-890-848-16  
; Sequence 16 Application US/09890848  
; Patent No. 6573053  
; GENERAL INFORMATION:  
; APPLICANT: Fitch, Greg  
; APPLICANT: Odeh, Rajesh Muru  
; TITLE OF INVENTION: Analysis Method  
; FILE REFERENCE: P89903  
; CURRENT APPLICATION NUMBER: US/09/890,848  
; PRIOR FILING DATE: 2001-08-03  
; PRIOR APPLICATION NUMBER: PCT/GB00/00346  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: BP 99300873.9  
; PRIOR FILING DATE: 1999-02-05

NUMBER OF SEQ ID NOS: 18  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 16  
LENGTH: 5277  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURES:  
OTHER INFORMATION: Description of Artificial Sequence: synthetic  
US-09-890-848-16

Query Match 67.4%; Score 2469.4; DB 4; Length 5277;  
Best Local Similarity 82.3%; Pred. No. 0;  
Matches 3014; Conservative 0; Mismatches 436; Indels 213; Gaps 7;

3 AATTGTAAAGCGTTAATATTTTGTAAATTCGCGTTAAATTTGTTAAATCGCTCAT 62  
1239 AAAATGTAAAGCGTTAATATTTTGTAAATTCGCGTTAAATTTTGTAAATCGCTCAT 1298  
63 TTTTAAACCAATAGCCGAAATCGCAAAATCCCTTAATAATCAAAAGATAGACGAGA 122  
1299 TTTTAAACCAATAGCCGAAATCGCAAAATCCCTTAATAATCAAAAGATAGACGAGA 1358  
123 TAGGTTAGAGTGTGTTCCAGTTTGGAAACAAGTCCACTATTAAGAAAGTGAATCCCA 182  
1359 TAGGTTAGAGTGTGTTCCAGTTTGGAAACAAGTCCACTATTAAGAAAGTGAATCCCA 1418  
183 ACGTCAAAAGGCGGAAAAACCGTCTATCAGGCGATGCGCCACTACGTGAACATCACTCC 242  
1419 ACGTCAAAAGGCGGAAAAACCGTCTATCAGGCGATGCGCCACTACGTGAACATCACTCC 1478  
243 AATCAAGTTTTTGGGGTGAAGTGCCTTAAGACTTAATCGAAACCTTAAGGAGACC 302  
1479 AATCAAGTTTTTGGGGTGAAGTGCCTTAAGACTTAATCGAAACCTTAAGGAGACC 1538  
303 CCCGATTAGAGCTTGAAGGAGAAACCGCGCAACGTGCGAGAAAGGAGGAGAAAG 362  
1539 CCCGATTAGAGCTTGAAGGAGAAACCGCGCAACGTGCGAGAAAGGAGGAGAAAG 1598  
363 CGAAAGAGGCGGCGCTAAGGCGCTGCAAGTGTAGCGGTCACTGCGGTGCTGCAACCA 422  
1599 CGAAAGAGGCGGCGCTAAGGCGCTGCAAGTGTAGCGGTCACTGCGGTGCTGCAACCA 1558  
423 CACCCGCGCGCTTAAATGCGCGCTACAGGCGCGCTC---CATTGCGCAATTCAGGCTGC 479  
1659 CACCCGCGCGCTTAAATGCGCGCTACAGGCGCGCTC---CATTGCGCAATTCAGGCTGC 1718  
480 GCAACTGTTGGGAAGGCGATGCTGCGGCGCTTGCCTAATTAACGAGCTGGGAAAG 539  
1719 GCAACTGTTGGGAAGGCGATGCTGCGGCGCTTGCCTAATTAACGAGCTGGGAAAG 1778  
540 GGGGATGTGCTGCAAGGCGATTAAGTGGGTAACGCGAGGTTTCCAGTCAAGCTT 599  
1779 GGGGATGTGCTGCAAGGCGATTAAGTGGGTAACGCGAGGTTTCCAGTCAAGCTT 1838  
600 GTAAACCAAGCGGCAAGGAGCGCGCTGCTTCAATTCAGCTTTTGAACCCGCGAGAGAG 659  
1839 GTAAACCAAGCGGCAAGGAGCGCGCTGCTTCAATTCAGCTTTTGAACCCGCGAGAGAG 1898  
660 GGGAGACTCGCGGTCGCAATGTGTTTACGCGTATGAGAGAGAGTGAAGTGTCCACA 719  
1899 AATTAAGATTTCAATTAAGTGTGTTTACGCGTATGAGAGAGAGTGAAGTGTCCACA 1958  
720 CGCTGCAAGACGAGCTAGATTAACCTTGAAGATATCATATGTGACGAGCTT 779  
1959 GGGCCCAACAAC-----TTCAATGAACGATCAATTAATTAACCACTAGAGAGAG 2010  
780 AAGATTAATGCGTAAATTTGACGATGATGATGATGATGATGATGATGATGATGATGAT 2011  
2011 AATCGATTAATTTTACCTTATTTCTGGGTAATTAATCAAGGAGAT---GATT 2066  
840 TTGGGTACGGGCGCCCTCGAGGTGACGCTTGCATTAAGTTGATTCGAATTTCTCG 899

2067 TTGATCTATTAACAGATATTAATGCAAAAACTGATTAACACTTTAATTAATCTTT 2126  
900 CAGCCGGGGGATTCATGATGTTCTAGACGCGCGCCACCGGGGAGCTCCAGCTTTTG 2127  
2127 CAGCATTTTCGGT-----TTGATTTAC 2148  
960 TTCCCTTTAGTAGGGTTAATTAATCCATGCGTCAATTTTACGAGACTATCTTTCTA 1019  
2149 TTCTATTCAATGTAATTAATTAAGTATCAATCAATCAATCAATCAATCAATCAATCA 2208  
1020 GGGTAAATCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2209  
2209 CAATTCATGATTAAGTACTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2268  
1080 CCAAGCTGGGCTATGCTGAGATCGGGAGGAAAGCCGTCCTTTTCCCGGAGGTT 1139  
2269 CGCC-----CTGATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2272  
1140 GAAAGCGCATGAAAGATTTGCCAGATGATGATGATGATGATGATGATGATGATGATGAT 2273  
2273 -----CTCCCGCCATCCGCTTAAACGAA 2300  
1200 CGACGTTTACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2301  
2301 GGAAGAGTTAGACAACTGAAAGTCTAGTCCCTATTTATTTTATTTATGTTATGTTAGT 2360  
1260 TTGTTCAAGGCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2361  
2361 ATTAAGAAAGTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 2400  
1320 GTACGCGGAGCGCATGACAAACCGGACCAATACCGGAGCGGAGCGGAGCGGAGCGG 2401  
2401 -----CTGATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2435  
1380 GGTGCGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2436  
2436 GAAACCTTCTTGAAGAGTTTGGAGCTCGAA----- 2471  
1440 CTTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2472  
2472 ---GGCTTTAATTTGCGGATCCAGCTTTTGTCTTGTAGTGGTAAATTCGAGCT 2528  
1500 TGGCGTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2529  
2529 TGGCGTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2588  
1560 ACAACATGAGAGCGGAGACATTAAGTAAAGCTGAGGTCCTATGATGATGATGATGAT 2589  
2589 ACAACATGAGAGCGGAGACATTAAGTAAAGCTGAGGTCCTATGATGATGATGATGATGAT 2648  
1620 TCACATTAATTTGCGTGGCTCACTGCGCTTTCAGTCTGAGGAAACCTGTCGACAGC 1679  
2649 TCACATTAATTTGCGTGGCTCACTGCGCTTTCAGTCTGAGGAAACCTGTCGACAGC 2708  
1680 TGCATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2709  
2709 TGCATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2768  
1740 CTTCTGCGCTCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2769  
2769 CTTCTGCGCTCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2828  
1800 ACTCAAGCGGTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2829  
2829 ACTCAAGCGGTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2888  
1860 GAGCAAAAGGCGGAGAAAGCGGAAACCTGTAATAAGCGGCTGATGATGATGATGATGAT 2889  
2889 GAGCAAAAGGCGGAGAAAGCGGAAACCTGTAATAAGCGGCTGATGATGATGATGATGATGAT 2948  
1920 ATAGGCTCGGCGCGCTGAGAGATCACAAAATGAGAGCTCAAGTCAAGGTTGCGAA 2949  
2949 ATAGGCTCGGCGCGCTGAGAGATCACAAAATGAGAGCTCAAGTCAAGGTTGCGAA 3008

QY 1980 ACCGACGAGACTATAAGATACAGAGGTTTCCCCCTGSAAGTCCCTCGTCCGCTC 2039  
DB ACCGACGAGACTATAAGATACAGAGGTTTCCCCCTGSAAGTCCCTCGTCCGCTC 3068  
QY 2040 CTGTTCCGACCCCTGCGCTTACCGGATACCTGTCCGCTTCTCCCTTCCGGAACGTGG 2099  
DB CTGTTCCGACCCCTGCGCTTACCGGATACCTGTCCGCTTCTCCCTTCCGGAACGTGG 3128  
QY 2100 CGCTTCTCATAGCTCAGCGTGTAGATCTCAAGTTCCGTTGAGTGGTCTTCCCAAC 2159  
DB CGCTTCTCATAGCTCAGCGTGTAGATCTCAAGTTCCGTTGAGTGGTCTTCCCAAC 3188  
QY 2160 TGGGCTGTGTCAGAGACCCCGCTTACGCGGACCGGCTTATCCGTTATCTAC 2219  
DB TGGGCTGTGTCAGAGACCCCGCTTACGCGGACCGGCTTATCCGTTATCTAC 3248  
QY 2220 GTCTTGAATCCAAACCCGTTAAGACAGACTTATCCGCACTGACAGACCACTGGTAA 2279  
DB GTCTTGAATCCAAACCCGTTAAGACAGACTTATCCGCACTGACAGACCACTGGTAA 3308  
QY 2280 GGATTTAGCAGACGAGATATGATGAGCGGTGCTACAGAGTTCTTGAAGTGTGGCTAA 2339  
DB GGATTTAGCAGACGAGATATGATGAGCGGTGCTACAGAGTTCTTGAAGTGTGGCTAA 3368  
QY 2340 ACGGCTACATAGAAAGGACAGATTTGGTATCTGGCTCTGTCAGAGCACTTACCTTCG 2399  
DB ACGGCTACATAGAAAGGACAGATTTGGTATCTGGCTCTGTCAGAGCACTTACCTTCG 3428  
QY 2400 GAAAAAGATGTAGCTCTTGATCCGCAAAACCAACGCGTGTAGCGGTGGTTTTT 2459  
DB GAAAAAGATGTAGCTCTTGATCCGCAAAACCAACGCGTGTAGCGGTGGTTTTT 3488  
QY 2460 TTGTTTGAACAGACAGATTAACGGCGCAAAAAAGATCTCAGAGAGATCTTGTATCT 2519  
DB TTGTTTGAACAGACAGATTAACGGCGCAAAAAAGATCTCAGAGAGATCTTGTATCT 3548  
QY 2520 TTTCTACGGGGCTGACGCTCAGTGAAGAAACCTCAGTTAAGGATTTTGGTCA 2579  
DB TTTCTACGGGGCTGACGCTCAGTGAAGAAACCTCAGTTAAGGATTTTGGTCA 3608  
QY 2580 GATTATCAAAAAGATCTTCACTAGATCTTTTAAATTAATAATGAAGTTTAAATCA 2639  
DB GATTATCAAAAAGATCTTCACTAGATCTTTTAAATTAATAATGAAGTTTAAATCA 3668  
QY 2640 TCTAAGTATATGATGATTAACCTTGTCTGACAGTTACCAATGCTTAACTAGAGCAC 2699  
DB TCTAAGTATATGATGATTAACCTTGTCTGACAGTTACCAATGCTTAACTAGAGCAC 3728  
QY 2700 CTATCTCAGGATCTGTCTATTTGGTTCATCTCATAGTTGCTGACTCCCGCTGTATGA 2759  
DB CTATCTCAGGATCTGTCTATTTGGTTCATCTCATAGTTGCTGACTCCCGCTGTATGA 3788  
QY 2760 TAACTACGATACGGAAGGCTTACATCTGCGCCAGTGTGCAATGATACCGGAGACC 2819  
DB TAACTACGATACGGAAGGCTTACATCTGCGCCAGTGTGCAATGATACCGGAGACC 3848  
QY 2820 CAGCTCAGCGGCTCCAGATTTATCAGATTAACAGCAGCCGGAAGGCGGAGCGCA 2879  
DB CAGCTCAGCGGCTCCAGATTTATCAGATTAACAGCAGCCGGAAGGCGGAGCGCA 3908  
QY 2880 GAAGTGTCTCCGCAACTTATCCGCTCCATCCAGTCTATTAATTTGGTCCGGAAGTA 2939  
DB GAAGTGTCTCCGCAACTTATCCGCTCCATCCAGTCTATTAATTTGGTCCGGAAGTA 3968  
QY 2940 GAGTAAATAGTTCCGCAAGTTATAGTTTGCAGACGTTTGCATTTCTACAGGATCG 2999  
DB GAGTAAATAGTTCCGCAAGTTATAGTTTGCAGACGTTTGCATTTCTACAGGATCG 4028  
QY 3000 TGGTGTCAAGCTGCTGTTTGTATGCTTCAATTCAGCTCCGGTTCCCAAGATCAAGGC 3059  
DB TGGTGTCAAGCTGCTGTTTGTATGCTTCAATTCAGCTCCGGTTCCCAAGATCAAGGC 4088

QY 3060 GAGTTACATGATCCCCCATGTTGTGCAAAAAAGGGTTAGCTCTTGGTCTCCGATCG 3119  
DB GAGTTACATGATCCCCCATGTTGTGCAAAAAAGGGTTAGCTCTTGGTCTCCGATCG 4148  
QY 3120 TTGTCAAGATAGTTGGCCGAGTGTATCATATGATTAAGCAGACATGCTAAAT 3179  
DB TTGTCAAGATAGTTGGCCGAGTGTATCATATGATTAAGCAGACATGCTAAAT 4208  
QY 3180 CTCTTACTGTATGATCCGTAAGATGCTTTCTGTGACGTGTGATGATCTCAACAGT 3239  
DB CTCTTACTGTATGATCCGTAAGATGCTTTCTGTGACGTGTGATGATCTCAACAGT 4268  
QY 3240 CATCTGAGATAGTATGAGGCGCAACGATGCTTGGCCGCGCTCAATACGGGATA 3299  
DB CATCTGAGATAGTATGAGGCGCAACGATGCTTGGCCGCGCTCAATACGGGATA 4328  
QY 3300 ATACCGCCCATATAGCAGAACTTTAAAGTCTCATCATTTGAAACGTTCTTGGGGC 3359  
DB ATACCGCCCATATAGCAGAACTTTAAAGTCTCATCATTTGAAACGTTCTTGGGGC 4388  
QY 3360 GAAAACTCTAAGATCTTACCGCTGTGAGATCAAGTGTGATCACTCGTGAC 3419  
DB GAAAACTCTAAGATCTTACCGCTGTGAGATCAAGTGTGATCACTCGTGAC 4448  
QY 3420 CCACTGATCTTCAAGATCTTTTACTTCAACAGCTTCTGGGTGAGCAAAACAGGAA 3479  
DB CCACTGATCTTCAAGATCTTTTACTTCAACAGCTTCTGGGTGAGCAAAACAGGAA 4508  
QY 3480 GGCAGAAATGCGGCAAAAAAGGAATTAAGGCGCACAGGAAATGTTGAATCTACTACT 3539  
DB GGCAGAAATGCGGCAAAAAAGGAATTAAGGCGCACAGGAAATGTTGAATCTACTACT 4568  
QY 3540 TCTTTTCAATATTAATTAAGCAATTAATCAAGGTTATGTTCTCATGAGCGGATCATAT 3599  
DB TCTTTTCAATATTAATTAAGCAATTAATCAAGGTTATGTTCTCATGAGCGGATCATAT 4628  
QY 3600 TTGAATGATTTAAGAAATTAACAATTAAGGCTTCCGCGCAATTTCCCGGAAAGTGC 3659  
DB TTGAATGATTTAAGAAATTAACAATTAAGGCTTCCGCGCAATTTCCCGGAAAGTGC 4688  
QY 3660 CAC 3662  
DB 4689 CAC 4691

RESULT 6  
US-09-402-266B-21  
Sequence 21, Application US/09402266B  
Patent No. 6537767  
GENERAL INFORMATION:  
APPLICANT: HINNEN, Albert  
APPLICANT: HEGEMANN, Johannes  
APPLICANT: MUNDER, Thomas  
APPLICANT: SCHUSTER, Tilmer  
APPLICANT: FELDMANN, Horst  
APPLICANT: KRAMER, Wilfried  
APPLICANT: ZIMMERMANN, Friedrich  
TITLE OF INVENTION: Process for Screening Antimycotically Active Substances  
FILE REFERENCE: 38005-0094  
CURRENT APPLICATION NUMBER: US/09/402,266B  
CURRENT FILING DATE: 1999-10-08  
PRIOR APPLICATION NUMBER: PCT/EP98/01904  
PRIOR FILING DATE: 1998-10-08  
PRIOR APPLICATION NUMBER: DE 1973572.2  
NUMBER OF SEQ ID NOS: 38  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 21  
LENGTH: 3956  
TYPE: DNA  
ORGANISM: Unknown  
FEATURE:

5' ATGTGAAGCTTATATTTTGGTAAATTCGCTTAAATTTTGGTAAATCGCTCATTT 64  
1 ATGTGAAGCTTATATTTTGGTAAATTCGCTTAAATTTTGGTAAATCGCTCATTT 64  
65 TTAAACCAATAGGCCGAAATCCGCAAAATCCCTTAATTAATCAAAAGAAATTAACGAGATA 124  
61 TTAAACCAATAGGCCGAAATCCGCAAAATCCCTTAATTAATCAAAAGAAATTAACGAGATA 124  
125 GGGTACAGTGTGTTTCAAGTTTGAACAAGAGTCCACTTTAAAGAACGTGACCTCCAAC 180  
121 GGGTGAAGTGTGTTTCAAGTTTGAACAAGAGTCCACTTTAAAGAACGTGACCTCCAAC 180  
185 GTCAAAAGGCGAAAAACCGCTCTATCAGGCGCATGAGCCCACTAGTAAACATCACTTAA 240  
181 GTCAAAAGGCGAAAAACCGCTCTATCAGGCGCATGAGCCCACTAGTAAACATCACTTAA 240  
245 TCAAGTTTTTTGGGCTGAGGTGCGCTAAAGCACTAAATGGAAACCTTAAAGGAGCCCC 304  
241 TCAAGTTTTTTGGGCTGAGGTGCGCTAAAGCACTAAATGGAAACCTTAAAGGAGCCCC 304  
305 CGAATTTAGACTTGAACGGGAAAGCGCGGCAACGTGCGAAGAAAGGAAGGAAGAACG 364  
301 CGAATTTAGACTTGAACGGGAAAGCGCGGCAACGTGCGAAGAAAGGAAGGAAGAACG 364  
365 AAAGAGCGGCGGCTAGGGCGCTGCGCAAGTGAAGCGTCAACGTCGCGCTAACCAACA 424  
361 AAAGAGCGGCGGCTAGGGCGCTGCGCAAGTGAAGCGTCAACGTCGCGCTAACCAACA 424  
425 CCGCGCGGCTTAATGCGCGCTACAGAGCGCGTCCCATTTGGCAATTCAGGCTGCGCAAC 480  
421 CCGCGCGGCTTAATGCGCGGCTACAGAGCGCGTCCCATTTGGCAATTCAGGCTGCGCAAC 480  
485 TGTGGGAAGGCGGATCGGTCGCGGCTCTTCCGCTAATGAGCAGCGTGGGAAAGGGGGA 544  
481 TGTGGGAAGGCGGATCGGTCGCGGCTCTTCCGCTAATGAGCAGCGTGGGAAAGGGGGA 544  
545 TGTGTGCAAGCGCAATTAAGTTGGGTAAAGCCAGGGTTTTCCAGTCAAGACGTTGTAA 604  
541 TGTGTGCAAGCGCAATTAAGTTGGGTAAAGCCAGGGTTTTCCAGTCAAGACGTTGTAA 604  
605 ACGAGCGCAGTAGAGCGCGCTCGTTCATTCAAGTTTTGAAACCCGTGAGGAAGGAGGAG 664  
601 ACGAGCGCAGTAGAGCGCGCTCGTTCATTCAAGTTTTGAAACCCGTGAGGAAGGAGGAG 664  
665 ACTGCGGCGTGAATGTGTTTTACAGCGGTAGGAGC----- 701  
661 GGTGGCGCTCTGAAGACTAGTGATCCGCACTGAGAGGCCAACCGCATAGATGCGCAAGT 720  
702 ----- 701  
721 GCACAAACAATTAATTAATACTACTAGTAATTAACCTAATTTCTTACATTTTGCAC 780  
702 ----- 701  
781 GAAATTTGCTATTTTGTTAAGTCTTTTACACCATTTGTCTCACACCTCCGCTTAATC 840  
744 AAACCTGAAAGATTAATCATATGTGTGACGTAAAGCTTAAAGTAATCATGTGCT 794  
841 AACCAATTAAGCATTTAATCTAAGCGATACCAACATTTTCTGGGTCATGTCAC 900  
795 -----AAATTTGACGATGGGATCTGTAATACGATCAATATAGGGGGAATTT--- 841  
901 AGCTAACATTAATTAATTAATTTTGGGTCCTCTTGGCTTCAACCAAGTGAAGAAATGCA 960

[illegible]





/ PRIOR FILING DATE: 1998-10-08  
 / PRIOR APPLICATION NUMBER: DE 19713572.2  
 / NUMBER OF SEQ ID NOS: 38  
 / SOFTWARE: PatentIn version 3.0  
 / SEQ ID NO 20  
 / LENGTH: 4102  
 / TYPE: DNA  
 / ORGANISM: Unknown  
 / FEATURE:  
 / NAME/KEY: misc feature  
 / LOCATION: (1..1)  
 / OTHER INFORMATION: Synthetic sequence  
 / US-09-402-266B-20

Query Match 62.9%; Score 2305.2; DB 4; Length 4102;  
 Best Local Similarity 76.9%; Pred. No. 0;  
 Matches 3150; Conservative 0; Mismatches 508; Indels 440; Gaps 10;

QY	5	ATTGTAAGCGTTAATATTTTGTAAATTCGGCTTAAATTTTGTAAATCAGCTCATTT	64
DB	1	ATTGTAAGCGTTAATATTTTGTAAATTCGGCTTAAATTTTGTAAATCAGCTCATTT	64
QY	65	TTTAAACCAATAGGCGCAAAATCGCAAAATCCCTTAAATCAAAAGAAATGACCGAGATA	124
DB	61	TTTAAACCAATAGGCGCAAAATCGCAAAATCCCTTAAATCAAAAGAAATGACCGAGATA	124
QY	125	GAGTTAGTGTGTTCCAGTTTGAACAAGTCCACTATTTAAAGAACTGGAATCGAAC	184
DB	121	GAGTTAGTGTGTTCCAGTTTGAACAAGTCCACTATTTAAAGAACTGGAATCGAAC	184
QY	185	GTCAAAAGGCGCAAAATCGCTTATCAGGCGCATGCGCCACTACGTAAACCATCCTTAA	244
DB	181	GTCAAAAGGCGCAAAATCGCTTATCAGGCGCATGCGCCACTACGTAAACCATCCTTAA	244
QY	245	TCAAGTTTTTGGGGGTGAGGTGCGCTTAAAGCACTAAATCGAAACCTTAAAGGAGCC	304
DB	241	TCAAGTTTTTGGGGGTGAGGTGCGCTTAAAGCACTAAATCGAAACCTTAAAGGAGCC	304
QY	305	CGATTTAGAGCTTGAACGGGAAACCGGCGAACTGCGCAAGAAAGGAAAGGAAAGCG	364
DB	301	CGATTTAGAGCTTGAACGGGAAACCGGCGAACTGCGCAAGAAAGGAAAGGAAAGCG	364
QY	365	AAAGAGAGCGGCGCTTAAAGGCGCTGCAAGTGAAGGCTGCTGCGCGCTGCAACACA	424
DB	361	AAAGAGAGCGGCGCTTAAAGGCGCTGCAAGTGAAGGCTGCTGCGCGCTGCAACACA	424
QY	425	CCGCGCGCTTAAATGCGCGCTTAAAGGCGCTGCAAGTGAAGGCTGCTGCGCGCAAC	484
DB	421	CCGCGCGCTTAAATGCGCGCTTAAAGGCGCTGCAAGTGAAGGCTGCTGCGCGCAAC	484
QY	485	TGTTGGAAAGGCGCATGCGTCCGCTTTCGCTAATTACGCGCACTGCGCAAGGCGGA	544
DB	481	TGTTGGAAAGGCGCATGCGTCCGCTTTCGCTAATTACGCGCACTGCGCAAGGCGGA	544
QY	545	TGTGTCGAAGCGCATTAAGTTGGGTAAAGCGGTTTTTCCAGTTCACGACGTTGTAA	604
DB	541	TGTGTCGAAGCGCATTAAGTTGGGTAAAGCGGTTTTTCCAGTTCACGACGTTGTAA	604
QY	605	ACGACGCGCGTGAAGCGCGCTGCTTCAATTCAGTTTGAACCGGTGAGAGAGCGGCG	664
DB	601	ACGACGCGCGTGAAGCGCGCTGCTTCAATTCAGTTTGAACCGGTGAGAGAGCGGCG	664
QY	665	ACTGCGCGTGAAGATGTTTAAACAGCGGATGAGAGAGAGAGAGAGAGAGAGAGAG	720
DB	661	ACTGCGCGTGAAGATGTTTAAACAGCGGATGAGAGAGAGAGAGAGAGAGAGAGAG	720
QY	711	TCGAATTCATTTTTTTTTTATCTTTTTTGAATTCGTTTCTTTGAAAATTTTTTG	780
DB	721	TCGAATTCATTTTTTTTTTATCTTTTTTGAATTCGTTTCTTTGAAAATTTTTTG	780
QY	711	----- 710	

DB	781	ATTGTAATCTCCGACAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG	840
QY	711	----- 710	
DB	841	ATTATACGATATGATGTTGAAGAAACATGAAATTCGACGATTTCTTAACCACTG	900
QY	711	---TGCTGACACGCTGACAGAAACAGCGCTGATGTTAACTCTG--- 751	
DB	901	ACAGAAACAAACATGACAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG	960
QY	752	----- 882	
DB	961	GTGCTGCTACTCATCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1020
QY	757	TATATCATTTTGA-----CGTACGTTAAAGTATATCATGCGTAAATTTG	801
DB	1021	AAACAACTGTGTGCTTCATTTGATGTTGTTACCAACAAAGAAATTTGAGTTGTTG	1080
QY	802	ACGATGAGATCTGTAATACACTCACTATGAGGCGAATGGTAAACCGGCGCCCTG	861
DB	1081	AACATTAAGTCCAAATTTGTTTAAACAAACATGATATCTGATTTT	1140
QY	862	AGGTGACGCTATCGATTAAC----- 882	
DB	1141	CAATGAGGCGCACTTAAGCGCTAAAGCATTTCCGAACTAATTTTTACTCT	1200
QY	883	-----TTGATATGAAATTCGACCGCGGAGAT	912
DB	1201	TCGAAGACAGAAATTTGCTGACATTTGTAATCACTCAATTTGCACTGCGGGTG	1260
QY	913	CCACTAGTTTGAAGCGCGCGC-----CACCGGTGAGACTCGACTT	956
DB	1261	TATACAAATGACAAATGGGCGACATTAACGAATGACACGCTGTGGGCGCTAGTA	1320
QY	957	TTGTTCCCTTTAATGAGG-----TTAA	979
DB	1321	TTGTTACCGGTTTGAAGCGCGCGCGCAAGAACTAAGAACTTAAGAGCGCTTTTGA	1380
QY	980	TTGATCCATGCGGTGCAATTTTACCACTATCTTTTCTAGGTTTAACTGATGATCA	1039
DB	1381	TGTTAGCAGAAATTTGTCAGAAAGGCTCTTATCTATCTGAGAAATTAATCAAGGTA	1440
QY	1040	GGATCATCTGCTGCG-----TCTTTTCCGCGCTCACTCATCGCCAGCTGCG	1089
DB	1441	TGCAATTTGCGAAGCGCAAAAGATTTGTTATCGCTTTATGCTCAAGAGACATG	1500
QY	1090	GCTATCTGCGCATCGGGAAGAAAGCCGCTTTTCCGCGAGTTGAAGCGCAT	1149
DB	1501	GTGAAAGAGATGAAGTTTACATTTGTTATATGACACCGGTGTGGGTTTAAATGCA	1560
QY	1150	GGAAAGATTTCCGAGATGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1209
DB	1561	AGGAGAGCGATTTGGGTCAACATATAGAACCTGATGATGATGATGATGATGATGAT	1260
QY	1210	CCATGATGATTTGGGAAGGTGTGCGCATTTGCAAAAGGAGATGATGATGATGATGAT	1320
DB	1621	ACATTTATTTTGTGGAAGAGCATTTTTCGCAAAAGGAGATGATGATGATGATGAT	1380
QY	1255	AACTGTTGTTGAGCGCACTGCGTACCAAGT-----CGTGGGCTTTTCCGGAACA	1309
DB	1681	AACTGTTGTTGAGCGCACTGCGTACCAAGT-----CGTGGGCTTTTCCGGAACA	1309
QY	1310	GTTCCGATGCTGACCGCGCAAGCCATCAGAACCCGAAATTAACCGGCGAC-----	1364
DB	1741	AACTGTTATTAAGTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1800
QY	1365	CGAACTGCGCTGCGGTGCGAGATTAATGACAGCGGTGCGCGCTGCGAGATTAAC	1424
DB	1801	TATATCAGTATTAACCAATTTCTCATGTTGCGCGCTGCGAGAAATGATGATGATGAT	1484
QY	1425	AGCGAGACCGGTATCTGCTGATGCGCGAGAAATGATGATGATGATGATGATGATGAT	1544
DB	1861	TATGATATCGTGAACCTGAGAGGCGCGCGGTACCAAGCTTTTGTGCTTTAGTGA	1920



QY 1485 CCCGCGGGGCGGCTTGGCGTAAATCAATAGCTGTTTCTGTGGAATTTTATC 1544  
DB 1921 GGTAAATTTGGCGCTTGGCGTAAATCAATAGCTGTTTCTGTGGAATTTTATC 1980  
QY 1545 CGCTCACAATTTCAACAACAATACGAGCGGGAAGATTAAGTAAAGCTTGGGGTGCCT 1604  
DB 1981 CGCTCACAATTTCAACAACAATACGAGCGGGAAGATTAAGTAAAGCTTGGGGTGCCT 2040  
QY 1605 AATGATGAGCTAACTACATTAATTTGCGTGCCTCACTGCGCTTTTCAATCGGGA 1664  
DB 2041 AATGATGAGCTAACTACATTAATTTGCGTGCCTCACTGCGCTTTTCAATCGGGA 2100  
QY 1665 ACCGTGCTGCGAGCTGATTAATTAATTCGCGCAACGCGCGGGAAGCGGTTTGGCTA 1724  
DB 2101 ACCGTGCTGCGAGCTGATTAATTAATTCGCGCAACGCGCGGGAAGCGGTTTGGCTA 2160  
QY 1725 TTGGGCGGCTCTTCCGCTTCTCGCTCACTGACTGCTGCGCTCGGTCGTTGCGCTGCGGC 1784  
DB 2161 TTGGGCGGCTCTTCCGCTTCTCGCTCACTGACTGCTGCGCTCGGTCGTTGCGCTGCGGC 2220  
QY 1785 GAGCGGATTCAGCTCACTCAAAAGCGGTAATACGTTATCCAAGAAATCAGGGGATTAACG 1844  
DB 2221 GAGCGGATTCAGCTCACTCAAAAGCGGTAATACGTTATCCAAGAAATCAGGGGATTAACG 2280  
QY 1845 CAGAAAGAATGAGTACGACAAAGCGCAAGAAAGCGGGAACCGTAAAGCGCGCT 1904  
DB 2281 CAGAAAGAATGAGTACGACAAAGCGCAAGAAAGCGGGAACCGTAAAGCGCGCT 2340  
QY 1905 TGCTGCGCTTTTTCATAGGCTCCGCGCTGACGAGCATCACAAAATGCAAGCTCA 1964  
DB 2341 TGCTGCGCTTTTTCATAGGCTCCGCGCTGACGAGCATCACAAAATGCAAGCTCA 2400  
QY 1965 GTCAAGGTGCGAAACCCGACAGGACTATAAAGTATCAAGCGTTTCCCTGGAACCT 2024  
DB 2401 GTCAAGGTGCGAAACCCGACAGGACTATAAAGTATCAAGCGTTTCCCTGGAACCT 2460  
QY 2025 CCCCTGTCGCTCTCTCTGTTCCGACCCCTGCGCTTAACCGGATACCTGTCCGCTTTCTCC 2084  
DB 2461 CCCCTGTCGCTCTCTCTGTTCCGACCCCTGCGCTTAACCGGATACCTGTCCGCTTTCTCC 2520  
QY 2085 CTTGCGGGAAGCGTGGCGCTTTCTCATAGCTCAACGCTGATGAGTATCTGAGTACG 2144  
DB 2521 CTTGCGGGAAGCGTGGCGCTTTCTCATAGCTCAACGCTGATGAGTATCTGAGTACG 2580  
QY 2145 TCGTTCCGTCGAAGCTGGCTGTGTGACGAACCCCGCTTCAAGCCGACCGCTGCGCT 2204  
DB 2581 TCGTTCCGTCGAAGCTGGCTGTGTGACGAACCCCGCTTCAAGCCGACCGCTGCGCT 2640  
QY 2205 TATCCGATTAATTCGTTTGAATGTCGAACCCGCTTAAGACAGACTTATCCGACTGGCAG 2264  
DB 2641 TATCCGATTAATTCGTTTGAATGTCGAACCCGCTTAAGACAGACTTATCCGACTGGCAG 2700  
QY 2265 CAGGCACTGATTAACGATTAAGACAGAGGATTAATGAGGCGGCTCAAGAGTTCTTGA 2324  
DB 2701 CAGGCACTGATTAACGATTAAGACAGAGGATTAATGAGGCGGCTCAAGAGTTCTTGA 2760  
QY 2325 AGTGTGGCTTAACGAGCTACCTAGAGAGACGATTTTGTATCTGCGCTGCTGCTGA 2384  
DB 2761 AGTGTGGCTTAACGAGCTACCTAGAGAGACGATTTTGTATCTGCGCTGCTGCTGA 2820  
QY 2385 AGCCAGTTAATCTTGGGAAAAAGAGTTGAGTCTTGAATCCGCAACAAACCGCTG 2444  
DB 2821 AGCCAGTTAATCTTGGGAAAAAGAGTTGAGTCTTGAATCCGCAACAAACCGCTG 2880  
QY 2445 GTACGCGTGTGTTTTTTTGTGCAAGACGAGATTAAGCGGAGAAAAAGATCTCAAG 2504  
DB 2881 GTACGCGTGTGTTTTTTTGTGCAAGACGAGATTAAGCGGAGAAAAAGATCTCAAG 2940  
QY 2505 AAGATCTTTGATCTTTTCTAAGGAGTGAACGCTCACTGGAACGAAATCTCAAGTTAAG 2564  
DB 2941 AAGATCTTTGATCTTTTCTAAGGAGTGAACGCTCACTGGAACGAAATCTCAAGTTAAG 3000

QY 2565 GGAATTTTGTATGAGATTAATCAAAAAGGATCTTCACTAGATCTTTTAAATTTAAAT 2624  
DB 3001 GGAATTTTGTATGAGATTAATCAAAAAGGATCTTCACTAGATCTTTTAAATTTAAAT 3060  
QY 2625 GAAATTTTAAATCAATCTAAAGTATTAATGAGTAACTTGGCTGACAGTTACCAATGCT 2684  
DB 3061 GAAATTTTAAATCAATCTAAAGTATTAATGAGTAACTTGGCTGACAGTTACCAATGCT 3120  
QY 2685 TAAATCAGTACGACACTTATCTCAGCGATCTGTCTAATTTGTTCAATCCATAGTTGCTGAC 2744  
DB 3121 TAAATCAGTACGACACTTATCTCAGCGATCTGTCTAATTTGTTCAATCCATAGTTGCTGAC 3180  
QY 2745 TCCCGGTGTGTATTAATCAATCAATGAGGAGGCTTACCAATCGGCGCGAGTGTGCA 2804  
DB 3181 TCCCGGTGTGTATTAATCAATCAATGAGGAGGCTTACCAATCGGCGCGAGTGTGCA 3240  
QY 2805 TGATACCGCGAGACCCAGCTCAACGCGCTCAGATTTATACGATTAACAGCAGCAGCG 2864  
DB 3241 TGATACCGCGAGACCCAGCTCAACGCGCTCAGATTTATACGATTAACAGCAGCAGCG 3300  
QY 2865 GAAAGCGCGAGCGAGAAATGCTCTGCAATTTATCCGCTCCATCCAGTCTAATTAAT 2924  
DB 3301 GAAAGCGCGAGCGAGAAATGCTCTGCAATTTATCCGCTCCATCCAGTCTAATTAAT 3360  
QY 2925 GTTCCCGGGAAGCTAGAGTATGAGTTCGCGAGTTAATGTTTGGCAAGCTGTGCGCA 2984  
DB 3361 GTTCCCGGGAAGCTAGAGTATGAGTTCGCGAGTTAATGTTTGGCAAGCTGTGCGCA 3420  
QY 2985 TTGCTACAGGATGTGGTGTCAAGCTGTCGTTGATGAGTTCATTTCACTCCGCTT 3044  
DB 3421 TTGCTACAGGATGTGGTGTCAAGCTGTCGTTGATGAGTTCATTTCACTCCGCTT 3480  
QY 3045 CCAACGATCAAGCGAGTATCATGATCCCCATGTTGTGCAAAAAGCGGTTAGCTCT 3104  
DB 3481 CCAACGATCAAGCGAGTATCATGATCCCCATGTTGTGCAAAAAGCGGTTAGCTCT 3540  
QY 3105 TCGGTCTCCGATGTTGTGAGAAATGTTGGCGGAGTGTATCACTCATGTTGTAAG 3164  
DB 3541 TCGGTCTCCGATGTTGTGAGAAATGTTGGCGGAGTGTATCACTCATGTTGTAAG 3600  
QY 3165 CAGACATGATTAATTTCTTACTGTCATGCAATCCGTAAGATCTTTTCTGACTGGTG 3224  
DB 3601 CAGACATGATTAATTTCTTACTGTCATGCAATCCGTAAGATCTTTTCTGACTGGTG 3660  
QY 3225 AGTACTCAACCAATGATCTTGAAGATAGTGTAGCGGCAACGAGTTGCTCTTGGCCGG 3284  
DB 3661 AGTACTCAACCAATGATCTTGAAGATAGTGTAGCGGCAACGAGTTGCTCTTGGCCGG 3720  
QY 3285 GGTCAATACGGGATTAATACCGCGCAATAGCAAACTTTTAAAGTGTCTCATTTGGAA 3344  
DB 3721 GGTCAATACGGGATTAATACCGCGCAATAGCAAACTTTTAAAGTGTCTCATTTGGAA 3780  
QY 3345 AACGTTCTTGGGCGGAAAACTCTCAAGGATCTTACCGCTGTGAGTTCAGTTGAT 3404  
DB 3781 AACGTTCTTGGGCGGAAAACTCTCAAGGATCTTACCGCTGTGAGTTCAGTTGAT 3840  
QY 3405 AACCACTGTGACCCCACTGATCTTCAAGATCTTTTCACTTCAACGAGGTTTCTGGGT 3464  
DB 3841 AACCACTGTGACCCCACTGATCTTCAAGATCTTTTCACTTCAACGAGGTTTCTGGGT 3900  
QY 3465 GAGCAAAAACGAGAGGCAAAATGCGCAAAAAGGAAATGAGGCGGACACGGAATGTT 3524  
DB 3901 GAGCAAAAACGAGAGGCAAAATGCGCAAAAAGGAAATGAGGCGGACACGGAATGTT 3960  
QY 3525 GAATACTCATATCTTCTTTTCAATTAATTAAGACATTTATCAGGTTATTTGTCTCA 3584  
DB 3961 GAATACTCATATCTTCTTTTCAATTAATTAAGACATTTATCAGGTTATTTGTCTCA 4040  
QY 3585 TGAGCGGATCATATTTGAATGTAATTAAGAAAAATAAACAATAGGGGTTTCCGCGCAT 3644  
DB 4021 TGAGCGGATCATATTTGAATGTAATTAAGAAAAATAAACAATAGGGGTTTCCGCGCAT 4080  
QY 3645 TTCCCGCAAAAGTCCAC 3662

Db

4081 TTCCCGAAAAGTCCAC 4098

RESULT 8  
US-09-402-266B-18  
Sequence 18, Application US/09402266B  
Patent No. 6537767  
GENERAL INFORMATION:  
APPLICANT: HINNEN, Albert  
APPLICANT: HEGEMANN, Johannes  
APPLICANT: MCNDER, Thomas  
APPLICANT: SCHUSTER, Tilmer  
APPLICANT: FELDMANN, Horst  
APPLICANT: KRAMER, Wilfried  
APPLICANT: ZIMMERMANN, Friedrich  
TITLE OF INVENTION: Process for Screening Anticancerically Active Substances  
FILE REFERENCE: 38005-0094  
CURRENT APPLICATION NUMBER: US/09/402,266B  
PRIOR FILING DATE: 1999-10-08  
PRIOR APPLICATION NUMBER: PCT/EP98/01904  
PRIOR FILING DATE: 1998-10-08  
PRIOR APPLICATION NUMBER: DE 19713572.2  
NUMBER OF SEQ ID NOS: 38  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 18  
LENGTH: 4088  
TYPE: DNA  
ORGANISM: Unknown  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: ( )..71  
OTHER INFORMATION: Synthetic sequence  
US-09-402-266B-18

Query Match 62.9%; Score 2302; DB 4; Length 4088;  
Best Local Similarity 76.1%; Pred. No. 0;  
Matches 3108; Conservative 0; Mismatches 550; Indels 426; Gaps 5;

QY 5 ATTGTAAAGCTTAATATTTTGTAAATTCGCGTTAAATTTTGTAAATTCAGTCAATT 64  
Db 1 ATTGTAAAGCTTAATATTTTGTAAATTCGCGTTAAATTTTGTAAATTCAGTCAATT 64  
QY 65 TTTAACCAATAGCGCAATCGCAAAATCCCTTAATATCAAAAGATAGACCGAGATA 124  
Db 61 TTTAACCAATAGCGCAAAATCGCAAAATCCCTTAATATCAAAAGATAGACCGAGATA 124  
QY 125 GGGTTGAGTGTGTCGATTGGAACAAGATCACTATTAAGAACTGAGTCCAAAC 184  
Db 121 GGGTTGAGTGTGTCGATTGGAACAAGATCACTATTAAGAACTGAGTCCAAAC 184  
QY 185 GTCAAGAGGCGAAAAACGCTCTATCAGGGGAGTCCCACTACGTAACCACTCCATA 244  
Db 181 GTCAAGAGGCGAAAAACGCTCTATCAGGGGAGTCCCACTACGTAACCACTCCATA 244  
QY 245 TCAAGTTTTTGGGGTGAAGTCCCTTAAGCACTAAATCGAACCTTAAGGGAGCC 304  
Db 241 TCAAGTTTTTGGGGTGAAGTCCCTTAAGCACTAAATCGAACCTTAAGGGAGCC 304  
QY 305 CGATTTAAGCTTGAAGCGGAAAGCGCGCAACGTCGCGGAAAGGAGGAGGAGGAGG 424  
Db 301 CGATTTAAGCTTGAAGCGGAAAGCGCGCAACGTCGCGGAAAGGAGGAGGAGGAGG 424  
QY 365 AAAAGAGCGGCGCTAGAGCGCTGCAAGTATAGCGTCAACGTCGCGGAAAGGAGG 484  
Db 361 AAAAGAGCGGCGCTAGAGCGCTGCAAGTATAGCGTCAACGTCGCGGAAAGGAGG 484  
QY 425 CCCGCGCGCTTAAGCGCGCTTACAGAGCGCGTCCATTGCGCATTGAGGCTGGCAAC 480  
Db 421 CCCGCGCGCTTAAGCGCGCTTACAGAGCGCGTCCATTGCGCATTGAGGCTGGCAAC 480

QY 485 TGTGGGAAGGCGATCGGTGCGGCGCTTCCGCTATTAACCCAGCTGGGAAAGGAGG 544  
Db 481 TGTGGGAAGGCGATCGGTGCGGCGCTTCCGCTATTAACCCAGCTGGGAAAGGAGG 544  
QY 545 TGTGCTGCAAGCGGATTAAGTTGGGTAAACGCAAGGTTTCCAGTCAAGAGTGTAA 604  
Db 541 TGTGCTGCAAGCGGATTAAGTTGGGTAAACGCAAGGTTTCCAGTCAAGAGTGTAA 604  
QY 605 ACCAGCGCGAGTGAAGCGCGCTGCTTCAATTCAGGTTTGAACCCGTGAGGAGCGGAG 664  
Db 601 ACCAGCGCGAGTGAAGCGCGCTGCTTCAATTCAGGTTTGAACCCGTGAGGAGCGGAG 664  
QY 665 ACTGCGGTGCAATGTGTTTACAGGAGTGAAG----- 700  
Db 661 GGTGCGCTCTAGAACTAGTGAATCCGCGCATAGGCGCTCGTTCAAGATGACAGTATA 720  
QY 701 --CAGATGAAGTGTGCAAGCTGCAAGCAACGCAAGCTAGATTAACTTAAAGATA 758  
Db 721 GAATGATGATTAACCTTGTATCTTCAAGTATCACTGTTGTATCATATTAAGTACA 780  
QY 759 ATGATATTTGAGCTACGTTAAAGATTAATCATAGGTTAAATTAAGCGATGGATCTGTA 818  
Db 781 TTGATAGTATCATATTAATCAATGATATATATATGATGATGATGATGATGATGAT 840  
QY 819 TACGACTCATATAGGCGGAATGGGTACCGGCGCGCGCGCGCGCGCGCGCGCGCG 878  
Db 841 GTGCTACTATTAAGAACCTTGTGAGGAGGAAATGTTGATGATGATGATGATGATGAT 900  
QY 879 AAGCTTGAATTCGAATTCGCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 927  
Db 901 GAGCTTCTTATAGCAACCGCAAGAGCGTTGAACGCACTTACATCGTATGATCAT 960  
QY 928 -----CG 980  
Db 961 CTGCGCTGCAAGCAATCAACGCGAGGATTTCTGTACCTCTGCAAGCTTCAAG 1020  
QY 981 TAGATCCATAGCTCAATTTTACGAGACTATCTTTCAGGGTTAACTAGCTGCAATCA 1040  
Db 1021 AAAATGCGGAGTATCTGCAAGAGAGATCTTCTTCTTCTTCTTCTTCTTCTTCT 1080  
QY 1041 GATCATATGTGCGGCTTTTTCGCGCTAGTATGCCCAAGCTGCGCTATCTGCGG 1100  
Db 1081 GACATGTGTAACCGCGCTGTTGAAAGATCAACCGCTGGAAGGCTCTATCA 1140  
QY 1101 ATCGGAGGAAAGCGCGCGCTTTT-----CGCGAGGTTGAAGCGGATGAAAGGT 1158  
Db 1141 AGCGCAATCTGATCCAAACCTTTTACTCAACGCGCAGTAGGCGCTTTAAAGC 1200  
QY 1159 TTGCGAGGATGACTGCTGCTGATTGA----- 1186  
Db 1201 TTGACGAGAGCAATCCGCAAGCTTCAAGTGTGATGATGATGATGATGATGATGAT 1260  
QY 1187 ----- 1186  
Db 1261 AATGCACTCAAGATTAGGACCAAGCGGAAATGCTTGGCAAGCATATATATGATG 1320  
QY 1187 CGTTGAGCGAAAGCGAAGCTTACCATATGATTTGCGAAGAGTGTGCGCATGACCTTT 1246  
Db 1321 CAGAAACCTTAATCTGTGTGAGATCACTTGAATGATGATGATGATGATGATGATGAT 1380  
QY 1247 TAAAGTGAATCTGTTGTTGAGCGCACTGAGATACCAATGATGATGATGATGATGAT 1396  
Db 1381 TGCTTCTCTCTTTTCTGGAAGATGAGTGTGATGATGATGATGATGATGATGATGAT 1440  
QY 1307 ACAAGTTCGAGTGTGAGCGCGCAAGGAGCATACAGCAACCCGAAATTAACCGGAG 1366  
Db 1441 AGAGATGCAATCTGAATCTTGTGTTTCAATTTGTAATGAGCTTACTAGGCTTTTCTCT 1500  
QY 1367 GAACTGCGCGCGCGGTGCAAT----- 1391  
Db 1501 TGCTATCTTGTCTTGTATCTTGTGCGCTGATTTTATGATATTTTGAAGAAAT 1560  
QY 1392 ----- 1391

Dp	1561	CACACTACTTTATATATATAGTATTAATCTATTTATGTGATATATCGGAATCGCTAAGAAAAA	1620
Qy	1392	-----	1391
Dp	1621	AAGAGTCATCCGCTAGCGGAAAAAAAATGAAAAATCATTTACGAGGCAATAAAAATA	1680
Qy	1392	-----	1391
Dp	1681	TAGAGTGTATAGAGGAGCGCAAGATATAGAAAAAGAAAAATTGCGGAAAGACCTGTG	1740
Qy	1392	-----	1391
Dp	1741	TTATGACTTCCCTGACTAATGCCGTGTTCAAAACAATACCTGGCAGTCACTCTTACGCTC	1800
Qy	1392	-----AATGACAGCGGTGCGCGCTGGGATATTTACGTACGAGACGAGAT	1438
Dp	1801	ACCAAGCTCTTAAAAACGCGCGCGCTGCGAGAAATTCATATACAAGCTTATCGATACCGTCG	1860
Qy	1439	TCCTGGCGTGATGCGCGCAAAATGGAACAATGATACCCCGTAGTTACCCCGCGCGCGC	1498
Dp	1861	ACCTCGAGGGGGGGCGCGGTACCCAGCTTTTGTTCCCTTATGAGGGTTAATTGCGCGC	1920
Qy	1499	TTGGCGTATATGATGATCATAGCTGTGTTTCTGTGAAAATTTGTATCCGCTCACAAATTC	1558
Dp	1921	TTGGCGTATATATGATGATCATAGCTGTGTTTCTGTGAAAATTTGTATCCGCTCACAAATTC	1980
Qy	1559	CACAAATATGAGCGCGGAGACATTAAGTATTAAGCTGGGGTGCCTAATGATGAGCTAA	1618
Dp	1981	CACAAATATGAGCGCGGAGACATTAAGTATTAAGCTGGGGTGCCTAATGATGAGCTAA	2040
Qy	1619	CTGACATTAATTTGGGTTGCGCTCACTGCGCGCTTCCAGTGGGGAACCTGTGCGGCGAG	1678
Dp	2041	CTGACATTAATTTGGGTTGCGCTCACTGCGCGCTTCCAGTGGGGAACCTGTGCGGCGAG	2100
Qy	1679	CTGCAATTAATGAAATCGGCGCAACGCGCGGGAGAGCGGTTTGCATTTGGGCGCTTCC	1738
Dp	2101	CTGCAATTAATGAAATCGGCGCAACGCGCGGGAGAGCGGTTTGCATTTGGGCGCTTCC	2160
Qy	1739	GCTTCCTCGCTCACTGACTCGCTGCGCTCGTGTGGGCTGCGCGGACGCGGTATCACT	1798
Dp	2161	GCTTCCTCGCTCACTGACTCGCTGCGCTCGTGTGGGCTGCGCGGACGCGGTATCACT	2220
Qy	1799	CACCTCAAAAGGCGGTATATCGGTTATCCAAATACAGGGGATTAACGAGAAACATG	1858
Dp	2221	CACCTCAAAAGGCGGTATATCGGTTATCCAAATACAGGGGATTAACGAGAAACATG	2280
Qy	1859	TGAGCAAAAGGCGAGCAAAAGGCCAGAAACGTTAAAAAGGCGCGGTGCTGGCCTTTTC	1918
Dp	2281	TGAGCAAAAGGCGAGCAAAAGGCCAGAAACGTTAAAAAGGCGCGGTGCTGGCCTTTTC	2340
Qy	1919	CATAGGCTCGCGCCCCCTGAGAGACATCAAAAATGACGCTCAAGTCAGAGGTGCGA	1978
Dp	2341	CATAGGCTCGCGCCCCCTGAGAGACATCAAAAATGACGCTCAAGTCAGAGGTGCGA	2400
Qy	1979	AAACCGACAGGACTATATAAGATACAGGCGTTTCCCTGGGAAGCTCCCTGTCGCGCT	2038
Dp	2401	AAACCGACAGGACTATATAAGATACAGGCGTTTCCCTGGGAAGCTCCCTGTCGCGCT	2460
Qy	2039	CTGTGTCGAGCCCTGCGGCTTACCGGATACCTGTGCGGCTTCTCCCTTGGGAAAGCGT	2098
Dp	2461	CTGTGTCGAGCCCTGCGGCTTACCGGATACCGGCTTCTCCCTTGGGAAAGCGT	2520
Qy	2099	GCGGTTTCTATACTACAGCTGTATAGATATCTCAGTTGCGGTATAGTGTCTGCTCAAG	2158
Dp	2521	GCGGTTTCTATACTACAGCTGTATAGATATCTCAGTTGCGGTATAGTGTCTGCTCAAG	2580
Qy	2159	CTGGGCTGTGTGACGAACCCCGCTTACGCGGACGCGTACGCGCTTATCGGATCAT	2218
Dp	2581	CTGGGCTGTGTGACGAACCCCGCTTACGCGGACGCGTATCGGATCAT	2640
Qy	2219	CGTCTTGAATGCAACCGGATAGACAGCACTTATGCGCACTGGCAGACGCCACTGTGTAAC	2278

Db	2641	CGTCTTAGAGTCCAAACCCGGTAAAGACAGACTTAATGCGCACTGCGACGACCACTGGTAAC	2700
Qy	2279	AGGATTACAGAGCGAGGTATGTAGCGCGTCTCAAGAGTTCTTGAATGCTGGCCCTAAC	2338
Db	2701	AGGATTAGCAGAGCGAGGTATGTAGCGCGTCTCAAGACTTCTTGAATGGTGGCCTTAAC	2760
Qy	2339	TACGGCTTCACTAAGAAAGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACTTTC	2398
Db	2761	TACGGCTTCACTAAGAAAGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACTTTC	2820
Qy	2339	GGAAAAAAGATTGGAGCTCTTGATTCGGGCAAAACAAACCGCTGGTAGCGGTGTTT	2458
Db	2821	GGAAAAAAGATTGGAGCTCTTGATTCGGGCAAAACAAACCGCTGGTAGCGGTGTTT	2880
Qy	2459	TTTGTGTTGCAGACGACAGATTACCGCGCAGAAAAAAGATCTCAAGAAAGATCCCTTGTATC	2518
Db	2881	TTTGTGTTGCAGACGACAGATTACCGCGCAGAAAAAAGATCTCAAGAAAGATCCCTTGTATC	2940
Qy	2519	TTTTCTACAGGGGCTCTACGCTCAAGTGAAGCAAACTCAGTTAAGGATTTTGGTCATG	2578
Db	2941	TTTTCTACAGGGGCTCTACGCTCAAGTGAAGCAAACTCAGTTAAGGATTTTGGTCATG	3000
Qy	2579	AGATTATCAAAAAGATCTTCCACTAGATCCTTTTAATTAAAAATGAAGTTTAAATCA	2638
Db	3001	AGATTATCAAAAAGATCTTCCACTAGATCCTTTTAATTAAAAATGAAGTTTAAATCA	3060
Qy	2639	ATCTAAAGTATATATGAGTAAACTTTGGCTGACAGATTACCAATGCTTAAATCAGTAGAGCA	2698
Db	3061	ATCTAAAGTATATATGAGTAAACTTTGGCTGACAGATTACCAATGCTTAAATCAGTAGAGCA	3120
Qy	2699	CCTAATCTCAGGGATCTGTCTATTTTCGTTCAATCCATAGATTGCTGACCTCCCGTGTGTAG	2758
Db	3121	CCTAATCTCAGGGATCTGTGTCTATTTTCGTTCAATCCATAGATTGCTGACCTCCCGTGTGTAG	3180
Qy	2759	ATTAACCTACGATACCGGAGGGCTTACATCTGCGCCCAAGTGTGATAGATTAACCGCGAGAC	2818
Db	3181	ATTAACCTACGATACCGGAGGGCTTACATCTGCGCCCAAGTGTGATAGATTAACCGCGAGAC	3240
Qy	2819	CCACGCTCAACCGGCTCCAGATTTTATTCAGCAATAAACACGACGACGCGGAGCGCGACGCG	2878
Db	3241	CCACGCTCAACCGGCTCCAGATTTTATTCAGCAATAAACACGACGCGGAGCGCGACGCG	3300
Qy	2879	AGAAATGGTCCCTGCACTTTATCCGCTCCATCCAGTCTATTAATTGTCGCGGAAGCT	2938
Db	3301	AGAAATGGTCCCTGCACTTTATCCGCTCCATCCAGTCTATTAATTGTCGCGGAAGCT	3360
Qy	2939	AGAGTAAAGTATGTTGCCAGTTAATATGTTGGCAAGTTGTTGCCATTGCTACAGCATC	2998
Db	3361	AGAGTAAAGTATGTTGCCAGTTAATATGTTGGCAAGTTGTTGCCATTGCTACAGCATC	3420
Qy	2999	GTCGTATCAGCGCTGTGTGGTTTGGATAGGCTTCAATCAGCTCCGGTTCGCAACGATCAAG	3058
Db	3421	GTCGTATCAGCGCTGTGTGGTTTGGATAGGCTTCAATCAGCTCCGGTTCGCAACGATCAAG	3480
Qy	3059	CGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGTTAGTCTCTCGTCCCGCATC	3118
Db	3481	CGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGTTAGTCTCTCGTCCCGCATC	3540
Qy	3119	GTTGTCAAGATTAAGTTGGCCGCGAGTGTATCACTCATAGGTTATGCGACATGCAATAT	3178
Db	3541	GTTGTCAAGATTAAGTTGGCCGCGAGTGTATCACTCATAGGTTATGCGACATGCAATAT	3600
Qy	3179	TCTCTTACCTGATGCAATCCGTAAGATGCTTTTCTGTGACTGTGATGATCTCAACCAAG	3238
Db	3601	TCTCTTACCTGATGCAATCCGTAAGATGCTTTTCTGTGACTGTGATGATCTCAACCAAG	3660
Qy	3239	TCATTTCTAGAAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGCGCTCAATACGGAGT	3298
Db	3661	TCATTTCTAGAAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGCGCTCAATACGGAGT	3720
Qy	3299	AATACCGGCGCATACACGAATCTTAAATGTCTCATCTTGAAGAAAGCTTCTTCGGGG	3358
Db	3721	AATACCGGCGCATACACGAATCTTAAATGTCTCATCTTGAAGAAAGCTTCTTCGGGG	3780

```

QY      3359  GAAAACCTCTCAAGAGATCTTACCGCTGTGTGAGATCCAGTTCATGTAATCCCACTCTGTGCA 3418
DB      3781  GGAABACTCTCAAGAGATCTTACCGCTGTGTGAGATCCAGTTCATGTAATCCCACTCTGTGCA 3840
QY      3419  CCCAAGCTGATCTTCAAGATCTTCTTACTTTCACACAGCGTTTCTGGGTGAGCAAAAACAGGA 3478
DB      3841  CCCAAGCTGATCTTCAAGATCTTCTTACTTTCACACAGCGTTTCTGGGTGAGCAAAAACAGGA 3900
QY      3479  AGCAAAATGCGCGCAAAAAGGAATTAAGGCGCACAGAAATGTTGAATATCTCATCTC 3538
DB      3901  AGCAAAATGCGCGCAAAAAGGAATTAAGGCGCACAGAAATGTTGAATATCTCATCTC 3590
QY      3539  TTCTTTTTCATATTTATTTGAAGCATTTATCAGGTTATTTGTCTCATGAGCGATACATA 4020
DB      3961  TTCTTTTTCATATTTATTTGAAGCATTTATCAGGTTATTTGTCTCATGAGCGATACATA 4020
QY      3599  TTGAATGATTTTGAATAAATAAACAATAGGGGTTCCCGGACATTTCCCGAAAGTG 3658
DB      4021  TTGAATGATTTTGAATAAATAAACAATAGGGGTTCCCGGACATTTCCCGAAAGTG 4080
QY      3659  CCAC 3662
DB      4081  CCAC 4084

```

## RESULT 9

```

US-09-993-170-1
; Sequence 1, Application US/09993170
; Patent No. 6743969
; GENERAL INFORMATION:
; APPLICANT: Valent, Barbara
; APPLICANT: Bryan, Gregory
; TITLE OF INVENTION: Modification of a Plant Disease Resistance Gene
; TITLE OF INVENTION: Specificity and Method for Engineering Altered
; FILE REFERENCE: BBI385 US NA
; CURRENT APPLICATION NUMBER: US/09/993,170
; PRIOR FILING DATE: 2001-11-26
; PRIOR FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 1
; LENGTH: 4119
; TYPE: DNA
; ORGANISM: Artificial Sequence
; OTHER INFORMATION: Description of Artificial Sequence: Plasmid
US-09-993-170-1

```

```

Query Match      62.2%; Score 2279.2; DB 4; Length 4119;
Best Local Similarity 76.6%; Pred. No. 0;
Matches 3151; Conservative 0; Mismatches 508; Indels 457; Gaps 11;

```

```

QY      1  CTAATTGTAAAGCGTTAATATTTTGTAAATTCGGCTTAAATTTTGTAAATCAGCTC 60
DB      4  CTAATTGTAAAGCGTTAATATTTTGTAAATTCGGCTTAAATTTTGTAAATCAGCTC 63
QY      61  ATTTTAAACCAATAGGCGGAAATGCGCAAAATCCCTATTAATCAAAAAGATGACCGA 120
DB      64  ATTTTAAACCAATAGGCGGAAATGCGCAAAATCCCTATTAATCAAAAAGATGACCGA 123
QY      121  GATAGGTTGAGTGTGTTCCAGTTTGGAACAAGATCCCACTAATTAAGAACTGGAGCTC 180
DB      124  GATAGGTTGAGTGTGTTCCAGTTTGGAACAAGATCCCACTAATTAAGAACTGGAGCTC 183
QY      181  CAACGTCAAAGGCGCAAAAACCGCTATCAGGCGGATGCGCCACTACGTAACATCAACC 240
DB      184  CAACGTCAAAGGCGCAAAAACCGCTATCAGGCGGATGCGCCACTACGTAACATCAACC 243
QY      241  CTAATCAGTTTTTTGGGTGAGGTGCGTAAAGACTAATCGAAACCTTAAGGGAG 300
DB      244  CTAATCAGTTTTTTGGGTGAGGTGCGTAAAGACTAATCGAAACCTTAAGGGAG 300

```

```

DB      244  CTAATCAGTTTTTTGGGTGAGGTGCGTAAAGACTAATCGAAACCTTAAGGGAG 303
QY      301  CCCCGATTTTAAAGCTTAAACGGGGAAGCCGGCAACATGTCGCAAAAGAGAGGA 360
DB      304  CCCCGATTTTAAAGCTTAAACGGGGAAGCCGGCAACATGTCGCAAAAGAGAGGA 363
QY      361  AGCAAAAGAGGCGGCGCTAGAGGCGCTGCAAGTGTAGCGTCAAGCTGCGTAAAC 420
DB      364  AGCAAAAGAGGCGGCGCTAGAGGCGCTGCAAGTGTAGCGTCAAGCTGCGTAAAC 423
QY      421  CACACCGCGCGGCTTAAATGCGCGTAAACAGGCGGCTCCATTCGCAATCAGCTGCG 480
DB      424  CACACCGCGCGGCTTAAATGCGCGTAAACAGGCGGCTCCATTCGCAATCAGCTGCG 483
QY      481  CAACCTTTGGAAGGCGATCCGTCGAGGCGCTCTTGGTATTAACGCAAGTGGCAAGG 540
DB      484  CAACCTTTGGAAGGCGATCCGTCGAGGCGCTCTTGGTATTAACGCAAGTGGCAAGG 543
QY      541  GGGATGTGTCGAAGGCGATTAAGTTGGGTAAACGAGGTTTCCAGTCAAGAGCTTG 600
DB      544  GGGATGTGTCGAAGGCGATTAAGTTGGGTAAACGAGGTTTCCAGTCAAGAGCTTG 603
QY      601  TAAACGAGCGCCAGTGA-----GCGGCGCTGTT 630
DB      604  TAAACGAGCGCCAGTGAATTTGTAATACATCACTATAGGCGAATTTGGTTACCGGCC 663
QY      631  CATTACGTTTTTGAACCCGTGAGAGAGCGGCAAGCTCGCGGTGCAAAATGTTTACAG 690
DB      664  CCCCCTCGAGGTGATCCGAGGTAAGGTCTAGAGGCGTGAATCTCACAATATTTTGTGA 723
QY      691  CGTATGAGAGATTAAGATGCTCGACAGCGTGCAGAACCGACGCT-----728
DB      724  ATTTCAAAATGCAAGGAGCATGAATATATGATATTAATTAATTTGGCCCACTTAC 783
QY      739  -----AATTAACTTGAAGAAATATCATTTTGTAGCTACGTTAAAGATTAATCAT 790
DB      784  CTTATATTAATTTGTGATGATATCTATTAATTAATTTGTAGCTTAAATTAAT 843
QY      791  GCGTAAATTTGAACGATGAGATCTGTATTAAGACTCACTA--TAGGCGGAATGGGTACC 848
DB      844  GCAGATTAATTTACATACGAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 903
QY      849  GGGCCCCCTCGAGGTGAGGCGGATGATGATGATGATGATGATGATGATGATGATGAT 908
DB      904  TATATTAACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 963
QY      909  GGATCAGTATTTGTAAGGCGGCGGATGATGATGATGATGATGATGATGATGATGATGAT 968
DB      964  TAAATTAATTTGCGTGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1023
QY      969  GTGAG-----GTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1023
DB      1024  TAGAGAAATTTGCTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1083
QY      1019  AGGTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1083
DB      1084  ACGATTTTGTGCAATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1047
QY      1048  -----TCGTCGGGTCTTTTTCGGCTCAGTATCGCCAAAGCTGATTTGGG 1099
DB      1144  TTAATTTTGCATCGATGCTTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1203
QY      1100  CAT-----CTTTTCAACCGTGTGATTTGGG 1102
DB      1204  CTTTCAACAATTTGGACCTTTTCAACCGGTTTACATTAATTAATTAATTAATTAATTAATTAAT 1263
QY      1103  -----CGGGAAGAGAGAGCCGTCCTT-----1127
DB      1264  ATATCAACGAGATTTTAAAGCGGCTTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1323
QY      1128  -----TCCCGGAGGTTGAAGCGGATGGAAGAGTATGTCGCTGCTGCA 1182
DB      1324  CGCAGGCTTCGAAATTCGTCGCGGCTTAAAGATTTGTCGAGCTGCTGCTGCTGCTGCTGCTGCT 1383

```

1183 TTGAGCTTGAGCGAAACGCA--CGTTACCATGATGATGGGAG----- 1227  
Db ATCAGCGCGTTAAAAATGACAATCGTTATTTAGATTAATCTTTAAACGACAGCAG 1443  
1228 ----- 1227  
1444 ATATTCAAACTGGGTTCAAAAAGATTTTAAAGAAATTTACAGGAATGTACAGGGACG 1503  
1228 -----GTGGCCATGCAAGCGCTTTAACGGTGAACGTGTGTTGAGGCCACT 1275  
1504 CGGAGAAATTTCTTAACCTGCGCAGATTAATAAGTTTATACGTCGCTCGAGAAAGAG 1563  
1276 GGGATCCAGTTGTCGGGCTTTTCCGACACAGTTCGAGTCCGATGTCACCCGAGCGCA 1335  
1564 TTCAATATTGGCGATGACCTTATTAAGAAAAAAATTTATATATGCTCTCTTCT 1623  
1336 TCAGCAACCCGAAACAAATACCGGCA----- 1362  
1624 TCACAAACCCGTTAAACAGCAGGAAATTAAGTCCGGTAAACAGATACAGTTATTAAC 1683  
1363 -----GCCGAACCTGCCGTGCCGTG 1383  
1684 ATGAAATGTCATATAATTTTAAAGAGTGAAGATTAATGTTACGAATGGGATGGGA 1743  
1384 TGCAGATTAAATGACAGCGGTGCGGCGCTGGGATATTAGTCAAGCAGAGACGGATCTG 1443  
1744 TTCAAAATTTGATATGATACAAAGATATTAACCCGACAGTTATCTATTTTTCAC 1803  
1444 GCTGATGTCGCGAAGATGACATGATATCCCGTGAATTAACCGGCGGG----- 1493  
1804 AATGTGACGTTATTAATATTGTTAAATACGTTATGTTGGAATGAGGAAATCGCGGT 1863  
1494 ----- 1493  
1864 CAATCAACATAGAGGATCCACTAGTTCTAGAGCGCGCCACCGCGGTGAGCTCCA 1923  
1494 -----CGCGTTGGCGTAAATCATGTGTCATAGCTGT 1523  
1924 GCTTTGTTCCCTTATGAGGGTTAATTTGAGGTTGGCGTAAATCATAGTTCATAGCTGT 1983  
1524 TTCCGTGTGTAATTTGTTATCCGCTCACAAATTCACACACATACAGACCGGAAAGCATTA 1583  
1984 TTCTGTGTGTAATTTGTTATCCGCTCACAAATTCACACACATACAGACCGGAAAGCATTA 2043  
1584 AGTGTAAAGCTGCGGTCCTTAATGATGAGCTAACTACATTAATTCGTTGCGCTCAC 1643  
2044 AGTGTAAAGCTGCGGTCCTTAATGATGAGCTAACTACATTAATTCGTTGCGCTCAC 2103  
1644 TGCCCGCTTTTCAAGTCGGGAAACCTGTGTGTCAGCTGCAATTAATGAAATGCGCAACGG 1703  
2104 TGCCCGCTTTTCAAGTCGGGAAACCTGTGTGTCAGCTGCAATTAATGAAATGCGCAACGG 2163  
1704 CGGGGAGAGCGGTTTGTGCTATTGAGGCGCTCTTCGCGCTTCTCGGCTCATAGCTCGCTGC 1763  
2164 CGGGGAGAGCGGTTTGTGCTATTGAGGCGCTCTTCGCGCTTCTCGGCTCATAGCTCGCTGC 2223  
1764 GCTCGGTGTTGCGCTGCGGAGCGGATACAGCTCATCAAGCGGTTAAATACGTTAT 1823  
2224 GCTCGGTGTTGCGCTGCGGAGCGGATACAGCTCATCAAGCGGTTAAATACGTTAT 2283  
1824 CCACAGAAATCAGGGGATTAACGACAGAAAGAAATGTGACAAAGGCGCAGAAAGGCCA 1883  
2284 CCACAGAAATCAGGGGATTAACGACAGAAAGAAATGTGACAAAGGCGCAGAAAGGCCA 2343  
1884 GGAACCGTTAAAAAGCGCGGTTGCTGGCGTTTTCATAGGCTCGCGCCCTGACGAGC 1943  
2344 GGAACCGTTAAAAAGCGCGGTTGCTGGCGTTTTCATAGGCTCGCGCCCTGACGAGC 2403  
1944 ATCACAAAATCGACGCTCAAGTCAAGAGTGGCGAAACCCGACAGACATTAAGATACC 2003  
2404 ATCACAAAATCGACGCTCAAGTCAAGAGTGGCGAAACCCGACAGACATTAAGATACC 2463

2004 AGGCGTTTCCCGCTGGAAGCTCCCTGAGGCTCTCCGTTTCGACCCCTGCGGTTACCG 2063  
2464 AGGCGTTTCCCGCTGGAAGCTCCCTGAGGCTCTCCGTTTCGACCCCTGCGGTTACCG 2523  
2064 GATACCTGTCCCGCTTCTCCCTTCGGGAAACGTGCGCTTCTCATAGCTCAGCGTTGA 2123  
2524 GATACCTGTCCCGCTTCTCCCTTCGGGAAACGTGCGCTTCTCATAGCTCAGCGTTGA 2583  
2124 GGTATCTCAGTTCCGTTGATGAGTGTGCTCTCAAGCTGGGCTGTGTGACCAACCCCGG 2183  
2584 GGTATCTCAGTTCCGTTGATGAGTGTGCTCTCAAGCTGGGCTGTGTGACCAACCCCGG 2643  
2184 TTCAGCCGACCGGTCGGCTTATCCGTTATCTGCTTGAATCCAAACCCGTTAAGAC 2243  
2244 ACAGATTATCGCAGCTGGCAGACCACTGGTAAACAGATTAAGAGACGAGATATGAG 2303  
2704 ACAGATTATCGCAGCTGGCAGACCACTGGTAAACAGATTAAGAGACGAGATATGAG 2763  
2304 GCGGTGCTACAGAGTTCTTGAAATGTTGCGCTTAACAGGCTACATGAAGACAGTAT 2363  
2764 GCGGTGCTACAGAGTTCTTGAAATGTTGCGCTTAACAGGCTACATGAAGACAGTAT 2823  
2364 TTGATATCTGCGCTCTGCTGAAGCAGTTACCTTCGAAAAAGATGTTGATCTTGAT 2423  
2824 TTGATATCTGCGCTCTGCTGAAGCAGTTACCTTCGAAAAAGATGTTGATCTTGAT 2883  
2424 CCGGCAAAACAAACCAACCGCTGCTGAGCGGTTTGTGTTGTAAGCAGAGATTACGC 2483  
2884 CCGGCAAAACAAACCAACCGCTGCTGAGCGGTTTGTGTTGTAAGCAGAGATTACGC 2943  
2484 GCAGAAAAAAGATCTCAAGAGATCTTTGATCTTTTCTACGCGGCTGACGCTGAT 2543  
2944 GCAGAAAAAAGATCTCAAGAGATCTTTGATCTTTTCTACGCGGCTGACGCTGAT 3003  
2544 GGAACGAAACCTCAGTTAAGGATTTTGTGTCATGAAATTAACAAAAAGATCTTCACT 2603  
3004 GGAACGAAACCTCAGTTAAGGATTTTGTGTCATGAAATTAACAAAAAGATCTTCACT 3063  
2604 AGATCTTTTAAATTAATAATGAATTTAAATCAATCTAAAGTATATATAGTAAACTT 2663  
3064 AGATCTTTTAAATTAATAATGAATTTAAATCAATCTAAAGTATATATAGTAAACTT 3123  
2664 GGTCTGACAGTTACCAATGCTTAATCAAGTGAAGCACTTATCTGACGATCTGTATTC 2723  
3124 GGTCTGACAGTTACCAATGCTTAATCAAGTGAAGCACTTATCTGACGATCTGTATTC 3183  
2724 GTTATCTCATAGTTGCTGACTCCCGCTCGTGTGATTAATCAATACAGTACGGAAGGCTTAC 2783  
3184 GTTATCTCATAGTTGCTGACTCCCGCTCGTGTGATTAATCAATACAGTACGGAAGGCTTAC 3243  
2784 CATCTGCCCCAGTGTGCAATGATACCGGAGACCCAGCTCAACGGCTCAGATTTAT 2843  
3244 CATCTGCCCCAGTGTGCAATGATACCGGAGACCCAGCTCAACGGCTCAGATTTAT 3303  
2844 CAGCAATTAACCAAGCCAGCGGAAGGCGCAGACGCAAGAGTGTCTGCAACTTATCCG 2903  
3304 CAGCAATTAACCAAGCCAGCGGAAGGCGCAGACGCAAGAGTGTCTGCAACTTATCCG 3363  
2904 CTTCCATCTCAGTCTAATTAATTTGTTGCGGGAAGCTAAGTAAGTATTCGCAATTA 2963  
3364 CTTCCATCTCAGTCTAATTAATTTGTTGCGGGAAGCTAAGTAAGTATTCGCAATTA 3423  
2964 GTTGGCAGCAAGTTGTTGCAATTCGTAAGGCAATCGTGTGTCAGCGCTGCTGTTGATA 3023  
3424 GTTGGCAGCAAGTTGTTGCAATTCGTAAGGCAATCGTGTGTCAGCGCTGCTGTTGATA 3483  
3024 TGCTTCAATTCAGCTCGGTTTCCCAACGATCAAGCGAGTTACATGATCCCCCATGTTGT 3083  
3484 TGCTTCAATTCAGCTCGGTTTCCCAACGATCAAGCGAGTTACATGATCCCCCATGTTGT 3543  
3084 GCAAAAAAGCGTTAGTCTCTGCGTCTCGATCGTTGTCAAGAAATGTTGCGCGCAG 3143



D	341	GAGTCGACCGGCATGCAAGCTTGCGCTGATCATGTGTCATGCTGTTTCCGTGTGAATT	3400
Q	1539	GTATTCGCTCACAAATTCACACAACTACGACCGGAGCATTAAGTGTAAAGCTGGG	1598
D	3401	GTATTCGCTCACAAATTCACACAACTACGACCGGAGCATTAAGTGTAAAGCTGGG	3460
Q	1599	GTGCTTAATGAGTGAAGTAACTGATTAATTCGTTGGGCTCACTGCCGCTTTCAGT	1658
D	3461	GTGCTTAATGAGTGAAGTAACTGATTAATTCGTTGGGCTCACTGCCGCTTTCAGT	3520
Q	1659	CGGAAAACCTGTGTCGACGCTGCAATTAATGATCGGCCACGCGCGGGAGAGCGGTT	1718
D	3521	CGGAAAACCTGTGTCGACGCTGCAATTAATGATCGGCCACGCGCGGGAGAGCGGTT	3580
Q	1719	TGCTATTGAGCGCTCTTCGCTTCTCGCTCACTGACTCGCTGAGCTTCGCTTCGCG	1778
D	3581	TGCTATTGAGCGCTCTTCGCTTCTCGCTCACTGACTCGCTGAGCTTCGCTTCGCG	3640
Q	1779	TGCGGAGACGGGTATCAGCTCATCAAAAGCGGTATACGGTTATCCAAGATCAGGG	1838
D	3641	TGCGGAGACGGGTATCAGCTCATCAAAAGCGGTATACGGTTATCCAAGATCAGGG	3700
Q	1839	ATAACGAGAAAGAACTGTGAGCAAAAGCGAGCAAAAGCCGATTAAGG	1898
D	3701	ATAACGAGAAAGAACTGTGAGCAAAAGCGAGCAAAAGCCGATTAAGG	3760
Q	1899	CGCGTTCCTGCGCTTTTTCATAGGCTCGCGCCCTGACGAGCATCAAAAATGAC	1958
D	3761	CGCGTTCCTGCGCTTTTTCATAGGCTCGCGCCCTGACGAGCATCAAAAATGAC	3820
Q	1959	GCTCAAGTCAGAGTGGCGAAACCCGACAGGATCTAATAATACAGGCTTTCCCTG	2018
D	3821	GCTCAAGTCAGAGTGGCGAAACCCGACAGGATCTAATAATACAGGCTTTCCCTG	3880
Q	2019	GAACTCCCTGCTGCTCTCTGTTCCGACCTGCGCTTACCGGATCTGTCGCGCT	2078
D	3881	GAACTCCCTGCTGCTCTCTGTTCCGACCTGCGCTTACCGGATCTGTCGCGCT	3940
Q	2079	TTTCCCTTGGGAAAGGTGGGCTTTCTCATAGTCAAGCTGTAGTATCTCAAGTTG	2138
D	3941	TTTCCCTTGGGAAAGGTGGGCTTTCTCATAGTCAAGCTGTAGTATCTCAAGTTG	4000
Q	2139	TGTAGTCTGCTCCAGCTGGGCTGTGTGACGAAACCCCGCTTACGCCGACCGT	2198
D	4001	TGTAGTCTGCTCCAGCTGGGCTGTGTGACGAAACCCCGCTTACGCCGACCGT	4060
Q	2199	GCCCTTATCCGGTATCTGTCTTGAATCAACCCGGTAAAGACGATTAATGCGAC	2258
D	4061	GCCCTTATCCGGTATCTGTCTTGAATCAACCCGGTAAAGACGATTAATGCGAC	4120
Q	2259	TGGCAGACCACTGTATACAGATTTAGAGAGGAGTATAGCGGTGTCTACAGT	2318
D	4121	TGGCAGACCACTGTATACAGATTTAGAGAGGAGTATAGCGGTGTCTACAGT	4180
Q	2319	TCTTGAAGGTGGCTTAACTACGGCTACCTAGAGGACAGTATTTGTATCTGCGCTC	2378
D	4181	TCTTGAAGGTGGCTTAACTACGGCTACCTAGAGGACAGTATTTGTATCTGCGCTC	4240
Q	2379	TGCTGAACCAAGTATCTTGGAAAAAGATTTGTAGTCTTGAATCCGCAACAAACA	2438
D	4241	TGCTGAACCAAGTATCTTGGAAAAAGATTTGTAGTCTTGAATCCGCAACAAACA	4300
Q	2439	CGCTGTAGCGGTGTTTTTTTGTGCAAGCAGAGATTACGCGGAAAAAAGAT	2498

Qy	CTCAAGAAAGATCCTTATCTCTTCTACGGGGCTGACGCTCAGTGGAAAGAAACTCAC	2549
Db	CTCAAGAAAGATCCTTATCTCTTCTCTCAAGGGGGCTGACGCTCAGTGGAAAGAAACTCAC	4361
Qy	GTTAAGGATTTTGGTCATGAGATTATCAAAAAGATCTTTCACCTTAGATCCTTTAAATT	2559
Db	GTTAAGGATTTTGGTCATGAGATTATCAAAAAGATCTTTCACCTTAGATCCTTTAAATT	4421





QY 2079 TTCTCCCTCGGGAAGCGTGCCTTTCTATAGCTCAGCTGATGATCTCACTTGG 2138  
 DB 3941 TTCTCCCTCGGGAAGCGTGCCTTTCTATAGCTCAGCTGATGATCTCACTTGG 4000  
 QY 2139 TGTAGTGTGTTGCTGCAAGCTGAGCTGATGAGCAAGCCCGCTTACGCCGACGCT 2198  
 DB 4001 TGTAGTGTGTTGCTGCAAGCTGAGCTGATGAGCAAGCCCGCTTACGCCGACGCT 4060  
 QY 2199 GCGCTTATCCGGGAATCTATGCTTGTGATCCAAACCGGTGATAGACAGCACTTATGCCAC 2258  
 DB 4061 GCGCTTATCCGGGAATCTATGCTTGTGATCCAAACCGGTGATAGACAGCACTTATGCCAC 4120  
 QY 2259 TGGCAGCAGCTCAGTGTGTAACAGGATTAAGCAGGAGTATGAGCGGTCTACAGAGT 2318  
 DB 4121 TGGCAGCAGCTCAGTGTGTAACAGGATTAAGCAGGAGTATGAGCGGTCTACAGAGT 4180  
 QY 2319 TCTTGAAGTGTGAGGCTTAACTACGCTACCTAGAGAGCAGTATTTGATCTGCGCTC 2378  
 DB 4181 TCTTGAAGTGTGAGGCTTAACTACGCTACCTAGAGAGCAGTATTTGATCTGCGCTC 4240  
 QY 2379 TGTGAAAGCACTTACCTTCCGAAAAAGTTGATGCTTGAATCCGCAAAACAACA 2438  
 DB 4241 TGTGAAAGCACTTACCTTCCGAAAAAGTTGATGCTTGAATCCGCAAAACAACA 4300  
 QY 2439 CCGCTGTAGCGGTGTTTGTGTTTGTGCAAGCAGATTAAGCGCAAAAAAGAT 2498  
 DB 4301 CCGCTGTAGCGGTGTTTGTGTTTGTGCAAGCAGATTAAGCGCAAAAAAGAT 4360  
 QY 2499 CTCAGAAAGATCTTTGATCTTTTCTACGCGGTCTGACGCTAGAGAAAGAACTCAG 2558  
 DB 4361 CTCAGAAAGATCTTTGATCTTTTCTACGCGGTCTGACGCTAGAGAAAGAACTCAG 4420  
 QY 2559 GTTAAAGGATTTGATGATGATTAATCAAAAAAGATCTTCACTAGATCTTTAAAT 2618  
 DB 4421 GTTAAAGGATTTGATGATGATTAATCAAAAAAGATCTTCACTAGATCTTTAAAT 4480  
 QY 2619 AAAAATGAATTTTAAATCAATCTAAAGTATATGAGTAAACTTGTGCTACAGTTACC 2678  
 DB 4481 AAAAATGAATTTTAAATCAATCTAAAGTATATGAGTAAACTTGTGCTACAGTTACC 4540  
 QY 2679 AATGCTTAATCAGTGAAGGACCTATCTCAGGATTTGCTATTTTCTGATCCATAGTTG 2738  
 DB 4541 AATGCTTAATCAGTGAAGGACCTATCTCAGGATTTGCTATTTTCTGATCCATAGTTG 4600  
 QY 2739 CCTGACTCCCGTCCGTGATGATTAATCAATCAAGTATCTGATTTTCTGATCCATAGTTG 2798  
 DB 4601 CCTGACTCCCGTCCGTGATGATTAATCAATCAAGTATCTGATTTTCTGATCCATAGTTG 4660  
 QY 2799 CTGCAATGATACCGGAGACCCAGCTCAGCGCTCCAGATTTATGAGCAATTAACCAAGC 2858  
 DB 4661 CTGCAATGATACCGGAGACCCAGCTCAGCGCTCCAGATTTATGAGCAATTAACCAAGC 4720  
 QY 2859 CAGCCGGAAGGGCGGAGAGGAGTGTCTGCAACTTATCCGCTCCATCCAGTCTA 2918  
 DB 4721 CAGCCGGAAGGGCGGAGAGGAGTGTCTGCAACTTATCCGCTCCATCCAGTCTA 4780  
 QY 2919 TTAATTTGTCGGGAAGTGAAGTGAAGTGTGCACTTAAATGTTTGGCAACGTTG 2978  
 DB 4781 TTAATTTGTCGGGAAGTGAAGTGAAGTGTGCACTTAAATGTTTGGCAACGTTG 4840  
 QY 2979 TTGCAATTTGATAGGATGTTGTTGATGATGATGATGATGATGATGATGATGATGATGAT 3038  
 DB 4841 TTGCAATTTGATAGGATGTTGTTGATGATGATGATGATGATGATGATGATGATGATGAT 3098  
 QY 3039 CCGGTTCCCAACGATCAAGGAGTGAATATCCCAATTTGCAAAAAAGCGGTTA 3098  
 DB 4901 CCGGTTCCCAACGATCAAGGAGTGAATATCCCAATTTGCAAAAAAGCGGTTA 4960  
 QY 3099 GCTCCTTGGTCTCGATCTGTTGTCAGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 3158  
 DB 4961 GCTCCTTGGTCTCGATCTGTTGTCAGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 5020

QY 3159 TTAATGAGCACTGATTAATCTCTTACTGATGATGATGATGATGATGATGATGATGATGAT 3218  
 DB 5021 TTAATGAGCACTGATTAATCTCTTACTGATGATGATGATGATGATGATGATGATGATGAT 5080  
 QY 3219 CTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3278  
 DB 5081 CTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5140  
 QY 3279 GCGCGGCTCAATCGGATTAATACCGGCACTAGAGCAAGTAAAGTCTTCACTCA 3338  
 DB 5141 GCGCGGCTCAATCGGATTAATACCGGCACTAGAGCAAGTAAAGTCTTCACTCA 5200  
 QY 3339 TTGAAAAAGTCTTCCGCGGCAAACTCTCAAGATCTTACCGTGTGATGATGATGATGAT 3398  
 DB 5201 TTGAAAAAGTCTTCCGCGGCAAACTCTCAAGATCTTACCGTGTGATGATGATGATGAT 5260  
 QY 3399 CGATGTAACCACTCTGCAACCACTGATCTTCAAGATCTTCAAGATCTTCAAGATCTTCAAG 3458  
 DB 5261 CGATGTAACCACTCTGCAACCACTGATCTTCAAGATCTTCAAGATCTTCAAGATCTTCAAG 5320  
 QY 3459 CTGGGTGAGCAAAACAGGAAAGCAAAATGCGCAAAAAAGGAAATAGGCGCAACGGA 3518  
 DB 5321 CTGGGTGAGCAAAACAGGAAAGCAAAATGCGCAAAAAAGGAAATAGGCGCAACGGA 5380  
 QY 3519 AATGTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3578  
 DB 5381 AATGTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5440  
 QY 3579 GTCTATAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5440  
 DB 5441 GTCTATAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5500  
 QY 3639 GCAATTTCCCGGAAGTCCAC 3662  
 DB 5501 GCAATTTCCCGGAAGTCCAC 5524

RESULT 12  
 US-09-573-322-20  
 / Sequence 20, Application US/09573322  
 / Patent No. 6531289  
 / GENERAL INFORMATION:  
 / APPLICANT: Bradley, John D.  
 / APPLICANT: Thompson, Craig M.  
 / APPLICANT: Moore, Jeffrey B.  
 / APPLICANT: Mobley, C. Richard  
 / APPLICANT: Bailey, David A.  
 / TITLE OF INVENTION: Regulated Gene Expression in Yeast and  
 / FILE REFERENCE: 0342/ID469-US4  
 / CURRENT APPLICATION NUMBER: US/09/573.322  
 / PRIOR APPLICATION NUMBER: 2000-05-18  
 / PRIOR FILING DATE: 1999-09-23  
 / PRIOR FILING DATE: 1998-08-21  
 / PRIOR APPLICATION NUMBER: 60/056,719  
 / NUMBER OF SEQ ID NOS: 30  
 / SOFTWARE: FASTSEQ for Windows Version 3.0  
 / SEQ ID NO: 20  
 / LENGTH: 7102  
 / TYPE: DNA  
 / ORGANISM: Artificial Sequence  
 / FEATURE:  
 / OTHER INFORMATION: ZM195 plasmid  
 US-09-573-322-20

Query Match 59.3%; Score 2172.8; DB 4; Length 7102;  
 Best Local Similarity 99.7%; Pred. No. 0;  
 Matches 2177; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
 QY 1479 GAGTTACCGCGGCGGCGCTTGGCTAATCATGATGATGATGATGATGATGATGATGATGAT 1538

Db 3341 GAGTCGACCGGCAATGCAAGCTTGGCTATATGCTCTATGCTGTTTCTGTGTAATT 3400  
Qy 1539 GTTATCCGCTCACAATTCACACAACATAGCAGCGGAGCAATAAGTGTAAAGCTGGG 1598  
Db 3401 GTTATCCGCTCACAATTCACACAACATAGCAGCGGAGCAATAAGTGTAAAGCTGGG 3460  
Qy 1599 GTGCTTAATGAGTACGTAACCTAATTAATTGGCTTGGCTCACTGGCCCTTTCCAGT 1658  
Db 3461 GTGCTTAATGAGTACGTAACCTAATTAATTGGCTTGGCTCACTGGCCCTTTCCAGT 3520  
Qy 1659 CGGGAACCTGTGTCAGCTGCACTTAATGAAATGCGCAACGCGCGGGAGAGCGGTT 1728  
Db 3521 CGGGAACCTGTGTCAGCTGCACTTAATGAAATGCGCAACGCGCGGGAGAGCGGTT 3580  
Qy 1719 TGCGTATTTGGGCGCTTTCCGCTTCTGCTCACTGACTGCTGCGCTCGGTCGTTGGC 1778  
Db 3581 TGCGTATTTGGGCGCTTTCCGCTTCTGCTCACTGACTGCTGCGCTCGGTCGTTGGC 3640  
Qy 1779 TGCGGCGAGCGGTAATCAGCTCACTCAAAAGCGGTAATCGGTTATCCACAGATTCAGGG 1838  
Db 3641 TGCGGCGAGCGGTAATCAGCTCACTCAAAAGCGGTAATCGGTTATCCACAGATTCAGGG 3700  
Qy 1839 ATAAAGCAGAAAGAAACATGTAGCAAAAGGCGCAAGAAAGGCGCAAGAAAGCGTAAAGG 1898  
Db 3701 ATAAAGCAGAAAGAAACATGTAGCAAAAGGCGCAAGAAAGGCGCAAGAAAGCGTAAAGG 3760  
Qy 1899 CCGGCTTGTGCGGCTTTTCCATAGAGCTCGCGCCCTGACGAGCATCAAAAAATCAC 1958  
Db 3761 CCGGCTTGTGCGGCTTTTCCATAGAGCTCGCGCCCTGACGAGCATCAAAAAATCAC 3820  
Qy 1959 GCTTAAGTCAGAGGTGGGGAACCCGACAGAGCTTAAGATACAGGCGTTTCCCTCG 2018  
Db 3821 GCTTAAGTCAGAGGTGGGGAACCCGACAGAGCTTAAGATACAGGCGTTTCCCTCG 3880  
Qy 2019 GAAGCTCCCTGTGCTGCTCTGTTCCGACCTGCGGCTTACCGGATACCTGTCCGCT 2078  
Db 3881 GAAGCTCCCTGTGCTGCTCTGTTCCGACCTGCGGCTTACCGGATACCTGTCCGCT 3940  
Qy 2079 TTCTCCCTTGGGAAGCGTGCGCTTCTCATAGCTCACTGTAAGTATCTCAGTTCCG 2138  
Db 3941 TTCTCCCTTGGGAAGCGTGCGCTTCTCATAGCTCACTGTAAGTATCTCAGTTCCG 4000  
Qy 2139 TGTAGTGTGTTGCTTCCAGCTGGGCTGTGCAAGAACCCCGCTTACGCCGCT 2198  
Db 4001 TGTAGTGTGTTGCTTCCAGCTGGGCTGTGCAAGAACCCCGCTTACGCCGCT 4060  
Qy 2199 GCGCTTATCCGCTTAATCTATGCTGAGTCAACCCGCTTAAGACAGCACTTATCGCAC 2258  
Db 4061 GCGCTTATCCGCTTAATCTATGCTGAGTCAACCCGCTTAAGACAGCACTTATCGCAC 4120  
Qy 2259 TGGGAGCAGCACTGTGTAACAGATTAAGCAGAGCGAGTATGAGCGGCTCTACAGAT 2318  
Db 4121 TGGGAGCAGCACTGTGTAACAGATTAAGCAGAGCGAGTATGAGCGGCTCTACAGAT 4180  
Qy 2319 TCTTGAAGTGTGGCTTAATCAAGGCTTACCTAAGAGACGATTTTGTGATCTGCGCTC 2378  
Db 4181 TCTTGAAGTGTGGCTTAATCAAGGCTTACCTAAGAGACGATTTTGTGATCTGCGCTC 4240  
Qy 2379 TGCTGAAGCAGTAACTTCCGAAAGAGTTGGAGCTTGTGATCCGGCGCAACAAACA 2438  
Db 4241 TGCTGAAGCAGTAACTTCCGAAAGAGTTGGAGCTTGTGATCCGGCGCAACAAACA 4300  
Qy 2439 CCGCTGTAGCGGTGTTTTTTTGTGTCAGAGCAGAGATTAACGCGCAGAAAAAGAT 2498  
Db 4301 CCGCTGTAGCGGTGTTTTTTTGTGTCAGAGCAGAGATTAACGCGCAGAAAAAGAT 4360  
Qy 2499 CTCAAGAAAGATCTTTTGAATCTTTTCTACGCGGCTGACGCTCAGTGTGAAACAAATCTAC 2558  
Db 4361 CTCAAGAAAGATCTTTTGAATCTTTTCTACGCGGCTGACGCTCAGTGTGAAACAAATCTAC 4420  
Qy 2559 GTTAAGGATTTTGTGATGATTAATCAAAAGAGATCTTCACTGATCTTTTAATT 2618

Db 4421 GTTAAGGATTTTGTGATGATTAATCAAAAGAGATCTTCACTGATCTTTTAATT 4480  
Qy 2619 AAAAATGAAGTTTAAATCAATCTTAAAGTATATGAGTAACTTGGTCTGACAGTTACC 2678  
Db 4481 AAAAATGAAGTTTAAATCAATCTTAAAGTATATGAGTAACTTGGTCTGACAGTTACC 4540  
Qy 2679 AATGCTTAATCAGAGGACCTATCTCAGGATCTGCTTAATTTGTTCAATCCADAGTGG 2738  
Db 4541 AATGCTTAATCAGAGGACCTATCTCAGGATCTGCTTAATTTGTTCAATCCADAGTGG 4600  
Qy 2739 CCTGACTCCCGTGTGATAGTAACTAGATAGTTCGCGAGGCGCTTACATCTGGCCCACTG 2798  
Db 4601 CCTGACTCCCGTGTGATAGTAACTAGATAGTTCGCGAGGCGCTTACATCTGGCCCACTG 4660  
Qy 2799 CTGCAATGATACCGCGAGACCCAGCGCTCACCGGCTCAAGATTTATCAGCAATAAACGAG 2858  
Db 4661 CTGCAATGATACCGCGAGACCCAGCGCTCACCGGCTCAAGATTTATCAGCAATAAACGAG 4720  
Qy 2859 CAGCGGAAGGCGGAGCGGAGAGTGTCTGCACTTTATCCGCTCCATCCAGTCTA 2918  
Db 4721 CAGCGGAAGGCGGAGCGGAGAGTGTCTGCACTTTATCCGCTCCATCCAGTCTA 4780  
Qy 2919 TTAATTTGCGCGGGAAGTATAGATAGTTCGCGAGTTAATAGTTTGGCAACGTTG 2978  
Db 4781 TTAATTTGCGCGGGAAGTATAGATAGTTCGCGAGTTAATAGTTTGGCAACGTTG 4840  
Qy 2979 TTGCAATGCTACAGGATCGTGTGTACAGCTGCTGCTTGTGATGGCTTCAATTCAGCT 3038  
Db 4841 TTGCAATGCTACAGGATCGTGTGTACAGCTGCTGCTTGTGATGGCTTCAATTCAGCT 4900  
Qy 3039 CCGGTTTCCCAACGATCAAGCGAGTTACATGATCCCGCATGTTGTGCAAAAAACGGTTA 3098  
Db 4901 CCGGTTTCCCAACGATCAAGCGAGTTACATGATCCCGCATGTTGTGCAAAAAACGGTTA 4960  
Qy 3099 GCTCCTTCGCTCCGATCGTGTGCAAGTATGTCGCGAGTGTATCACTCAGTG 3158  
Db 4961 GCTCCTTCGCTCCGATCGTGTGCAAGTATGTCGCGAGTGTATCACTCAGTG 5020  
Qy 3159 TTATGGAGCAGCTGATTAATTTCTTATCTGTATGCCATCCGTAGAGTCTTTCTGTGA 3218  
Db 5021 TTATGGAGCAGCTGATTAATTTCTTATCTGTATGCCATCCGTAGAGTCTTTCTGTGA 5080  
Qy 3219 CTGCTGATGATCAACCAATCTTCTGAGATATGTTAGCGGCGACCCAGTGTCTT 3278  
Db 5081 CTGCTGATGATCAACCAATCTTCTGAGATATGTTAGCGGCGACCCAGTGTCTT 5140  
Qy 3279 GCCCGGCTCAATCCGGAATTAATCCGCGCACATAGCAGAACTTTAAAGTGTCTATCA 3338  
Db 5141 GCCCGGCTCAATCCGGAATTAATCCGCGCACATAGCAGAACTTTAAAGTGTCTATCA 5200  
Qy 3339 TTGGAAGAGTCTTCTGCGGCGAAGAACTCTCAAGGATCTTACCGCTGTGAGATCCAGT 3398  
Db 5201 TTGGAAGAGTCTTCTGCGGCGAAGAACTCTCAAGGATCTTACCGCTGTGAGATCCAGT 5260  
Qy 3399 CGATGTAACCACTCGTGACCCCACTGATCTTCAAGATCTTTTACCAAGCGTTT 3458  
Db 5261 CGATGTAACCACTCGTGACCCCACTGATCTTCAAGATCTTTTACCAAGCGTTT 5320  
Qy 3459 CTGGGTAGGAAAAAGAGGCAAAATGCGCGCAAAAGGGAATAAGGCGCACAGGA 3518  
Db 5321 CTGGGTAGGAAAAAGAGGCAAAATGCGCGCAAAAGGGAATAAGGCGCACAGGA 5380  
Qy 3519 AATGTTGAATACTATATCTTCTTTTCAATATTAATGAAGATTTATCAAGGTTAAT 3578  
Db 5381 AATGTTGAATACTATATCTTCTTTTCAATATTAATGAAGATTTATCAAGGTTAAT 5440  
Qy 3579 GTCTCAGAGCGGATATATTTGATATGTTTGAAGAAATAAACAATAGGGGTTCCG 3638  
Db 5441 GTCTCAGAGCGGATATATTTGATATGTTTGAAGAAATAAACAATAGGGGTTCCG 5500  
Qy 3639 GCACATTTCCCGAAAAAGTGCAC 3662  
Db 5501 GCACATTTCCCGAAAAAGTGCAC 5524



Db 5256 TTATGCGAGCACTGATCAATTTCTTACTGTCATCCGTAAGATGCTTTTCTGTGA 5315  
Qy 3219 CTGTGAGTACTCAACCAAGTCATTTCTGAGAAATAGTATGCGGACCGAGTTCCTCTT 3278  
Db 5316 CTGTGAGTACTCAACCAAGTCATTTCTGAGAAATAGTATGCGGACCGAGTTCCTCTT 5375  
Qy 3279 GCGCGGCGTCAATACGGGATTAACCGGCGCACTAGAGAACTTTAAAGTCTCATCA 3338  
Db 5376 GCGCGGCGTCAATACGGGATTAATACCGGCGCACTAGAGAACTTTAAAGTCTCATCA 5435  
Qy 3339 TTGGAACAGTCTTCGCGGCGAAACCTCAAGGATCTTACCGCTGTGAGATCCAGTT 3398  
Db 5436 TTGGAACAGTCTTCGCGGCGAAACCTCAAGGATCTTACCGCTGTGAGATCCAGTT 5495  
Qy 3399 CGATGTAAACCACTGTCGACCCCACTGATCTTCAGCATCTTTTACCTTCACACGCTTT 3458  
Db 5496 CGATGTAAACCACTGTCGACCCCACTGATCTTCAGCATCTTTTACCTTCACACGCTTT 5555  
Qy 3459 CTGGGTGAGCAAAAACAGAAAGCAAAATGCGCAAAAAGGGAATTAAGGCGACACGGA 3518  
Db 5556 CTGGGTGAGCAAAAACAGAAAGCAAAATGCGCAAAAAGGGAATTAAGGCGACACGGA 5615  
Qy 3519 AATGTGAATCTCATCTCTCTCTTTTCAATATTTTGAAGCATTTATCAGGGTTATT 3578  
Db 5616 AATGTGAATCTCATCTCTCTCTTTTCAATATTTTGAAGCATTTATCAGGGTTATT 5675  
Qy 3579 GTCTCATGAGCGGATACATATTTGAATGATTTAGAAAATTAACAAATAGGCGTTCCGC 3638  
Db 5676 GTCTCATGAGCGGATACATATTTGAATGATTTAGAAAATTAACAAATAGGCGTTCCGC 5735  
Qy 3639 GCACATTTCCCGGAAAGTGCCAC 3662  
Db 5736 GCACATTTCCCGGAAAGTGCCAC 5759

RESULT 14  
US-09-404-066-21  
Sequence 21, Application US/09404066  
Patent No. 6365409  
GENERAL INFORMATION:  
APPLICANT: Bradley, John D.  
APPLICANT: Thompson, Craig M.  
APPLICANT: Moore, Jeffrey B.  
APPLICANT: Mobbe, C. Richard  
APPLICANT: Healy, Judith M.  
FILE OF INVENTION: REGULATED GENE EXPRESSION IN YEAST  
FILE REFERENCE: 0342/1D469US1  
CURRENT APPLICATION NUMBER: US/09/404,066  
PRIOR FILING DATE: 1999-09-23  
PRIOR APPLICATION NUMBER: US 09/138,024  
PRIOR FILING DATE: 1998-08-21  
PRIOR APPLICATION NUMBER: 60/056,719  
PRIOR FILING DATE: 1997-08-22  
NUMBER OF SEQ ID NOS: 24  
SOFTWARE: FastSeq For Windows Version 3.0  
SEQ ID NO 21  
LENGTH: 7333  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Plasmid pZM197  
US-09-404-066-21

Query Match 59.3%; Score 2172.8; DB 3; Length 7333;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 2177; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 1479 GAGTTACCGGCGCGGCGGCTTGGCGTATCATGTCTATCTGTTCTGTGTAATT 1538  
Db 3576 GAGTCGACCGGCGGCGGCGGCTTGGCGTATCATGTCTATCTGTTCTGTGTAATT 3635

Qy 1539 GTTATCCGCTCACAATTCCACAACATAGAGCGGAAAGCATTAAGTATAAGCGCTGG 1598  
Db 3636 GTTATCCGCTCACAATTCCACAACATAGAGCGGAAAGCATTAAGTATAAGCGCTGG 3695  
Qy 1599 GTGCTTAATGAGTGAAGTCACTCAATTAATGCGTTGCGCTCACTGCGCTTTTCCAGT 1658  
Db 3696 GTGCTTAATGAGTGAAGTCACTCAATTAATGCGTTGCGCTCACTGCGCTTTTCCAGT 3755  
Qy 1659 CCGGAACCTGTCTGCGCAGTGCATTAATGATTCGCAACCGCGGGAAGAGCGGTT 1718  
Db 3756 CCGGAACCTGTCTGCGCAGTGCATTAATGATTCGCAACCGCGGGAAGAGCGGTT 3815  
Qy 1719 TCGGATTTGGGCGCTCTTCGCTCTCTGCTCACTGACTCGCTGCGTCTGCTTGGC 1778  
Db 3816 TCGGATTTGGGCGCTCTTCGCTCTCTGCTCACTGACTCGCTGCGTCTGCTTGGC 3875  
Qy 1779 TCGGCGAGCGGTATCAGCTCACTCAAGCGGTATATCGGTTATCCACAGAAATCGAGG 1838  
Db 3876 TCGGCGAGCGGTATCAGCTCACTCAAGCGGTATATCGGTTATCCACAGAAATCGAGG 3935  
Qy 1839 ATTAACGAGAAAGAACATGTGACCAAAAGGCGCAAAAGGCGCAAAAGCGTAAAAAG 1898  
Db 3936 ATTAACGAGAAAGAACATGTGACCAAAAGGCGCAAAAGGCGCAAAAGCGTAAAAAG 3995  
Qy 1899 CCGGCTGCTGCGCTTTTCCATAGCGCTCGCGCGCGCGCGCGCGCGCGCGCGCGCT 1958  
Db 3996 CCGGCTGCTGCGCTTTTCCATAGCGCTCGCGCGCGCGCGCGCGCGCGCGCGCGCT 4055  
Qy 1959 GCTCAAGTCAAGGTGCGGAACCGGACAGACTATAAGATATACAGGCTTTCCCGCTG 2018  
Db 4056 GCTCAAGTCAAGGTGCGGAACCGGACAGACTATAAGATATACAGGCTTTCCCGCTG 4115  
Qy 2019 GAAGCTCCCTCTGTCGCTCTCTGTTCCGACCTTCGCGCTTACCGGATACCTGTCGCT 2078  
Db 4116 GAAGCTCCCTCTGTCGCTCTCTGTTCCGACCTTCGCGCTTACCGGATACCTGTCGCT 4175  
Qy 2079 TTCTCCCTTGGGGAAGGTGCGGCTTCTCACTACGCGTATAGATATCTGAGTCCG 2138  
Db 4176 TTCTCCCTTGGGGAAGGTGCGGCTTCTCACTACGCGTATAGATATCTGAGTCCG 4235  
Qy 2139 TGTAGTGTCTGCTCCGCTCAAGCTGCGCTGTGACGAAACCCCGCTTACGCGACGCT 2198  
Db 4236 TGTAGTGTCTGCTCCGCTCAAGCTGCGCTGTGACGAAACCCCGCTTACGCGACGCT 4295  
Qy 2199 GCGCTTATCCGCTATCTATGCTTGAAGTCAACCGGTAAGACAGACTTATCGCCAC 2258  
Db 4296 GCGCTTATCCGCTATCTATGCTTGAAGTCAACCGGTAAGACAGACTTATCGCCAC 4355  
Qy 2259 TGGACAGACCACTGTGTAACGAAATTAAGCAGAGGATATGAGCGGCTCTACAGAT 2318  
Db 4356 TGGACAGACCACTGTGTAACGAAATTAAGCAGAGGATATGAGCGGCTCTACAGAT 4415  
Qy 2319 TCTTGAATGTGTGCTTAACTACGAGCTACATGAAGACAGATTTTGGTATCGCGCTC 2378  
Db 4416 TCTTGAATGTGTGCTTAACTACGAGCTACATGAAGACAGATTTTGGTATCGCGCTC 4475  
Qy 2439 TGTGAAGCCAGTTACTCTGGAAGAAAGATGTGTAGCTCTTGAATCGGCAACAAACCA 2438  
Db 4476 TGTGAAGCCAGTTACTCTGGAAGAAAGATGTGTAGCTCTTGAATCGGCAACAAACCA 4535  
Qy 2439 CCGCTGTGAGCGGCTTTTGTGTTGCAAGCAGAGATTAACCGGCAAGAAAAAGAT 2498  
Db 4536 CCGCTGTGAGCGGCTTTTGTGTTGCAAGCAGAGATTAACCGGCAAGAAAAAGAT 4595  
Qy 2499 CTCAAGAAAGATCTTTGATCTTTCTCAAGGCGTCAAGCTCAGTGGACGAAGAACTCAC 2558  
Db 4596 CTCAAGAAAGATCTTTGATCTTTCTCAAGGCGTCAAGCTCAGTGGACGAAGAACTCAC 4655  
Qy 2559 GTTAAGGATTTTGTGATGATTAATCAAAAGATCTTCACTAGATCTTTTAATTT 2618  
Db 4656 GTTAAGGATTTTGTGATGATTAATCAAAAGATCTTCACTAGATCTTTTAATTT 4715  
Qy 2619 AAAATAGATTTTAATCAATCTAAAGTATATATGATTAACCTTGGTCTGACAGTTACC 2678



4716 AAAAAAGTTTAAATCAATTAAGATATATAGAACTTGGCTACAGCTTACC 4775  
4775 AATGCTTAATAGTAGAGCACTATCTAGCATCTGCTATTTCTGCTATCCATAGTTG 2738  
4776 AATGCTTAATAGTAGAGCACTATCTAGCATCTGCTATTTCTGCTATCCATAGTTG 2738  
2739 CCTGATCCCGGCGGATGATTAATAGATACGGAGAGGCTTACATCTGGCCCCAGTG 4835  
4836 CCGACTCCCGGCGGATGATTAATAGATACGGAGAGGCTTACATCTGGCCCCAGTG 4835  
2799 CTGCAATGATACCGGAGAACCCACGCTCAGCGCTCCAGATTTATGCAATTAACACG 2858  
4896 CTGCAATGATACCGGAGAACCCACGCTCAGCGCTCCAGATTTATGCAATTAACACG 4955  
2859 CAGCGGAGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2918  
4956 CAGCGGAGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 5015  
2919 TTAATTTGTCGGGAGAGTGAATAGTATGCTGAGTATGCTGAGTATGCTGAGTATG 2978  
5016 TTAATTTGTCGGGAGAGTGAATAGTATGCTGAGTATGCTGAGTATGCTGAGTATG 5075  
2979 TTGCAATTTGTCGGGAGAGTGAATAGTATGCTGAGTATGCTGAGTATGCTGAGTATG 3038  
5076 TTGCAATTTGTCGGGAGAGTGAATAGTATGCTGAGTATGCTGAGTATGCTGAGTATG 5135  
3039 CCGGTTCCCAACGATCAAGGCGAGTACATGATCCCGGAGTGGTGGCAAAAAGCGTTA 3098  
5136 CCGGTTCCCAACGATCAAGGCGAGTACATGATCCCGGAGTGGTGGCAAAAAGCGTTA 5195  
3099 GCTCCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCC 3158  
5196 GCTCCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCC 3278  
3159 TTATGCGACACTGATTAATCTCTTAATCTGATGCTGATGCTGATGCTGATGCTGATG 3218  
5256 TTATGCGACACTGATTAATCTCTTAATCTGATGCTGATGCTGATGCTGATGCTGATG 5315  
3219 CTGATGATGATCAACCAAGTATCTGATGATGATGATGATGATGATGATGATGATGATG 5315  
5316 CTGATGATGATCAACCAAGTATCTGATGATGATGATGATGATGATGATGATGATGATG 5375  
3279 GCGCGGCGTCAATACGAGGATTAATACGCGGCGATGATGATGATGATGATGATGATG 5375  
5376 GCGCGGCGTCAATACGAGGATTAATACGCGGCGATGATGATGATGATGATGATGATG 5435  
3339 TTGAAAAAGTTTAAATCAATTAAGATATATAGAACTTGGCTACAGCTTACC 3398  
5436 TTGAAAAAGTTTAAATCAATTAAGATATATAGAACTTGGCTACAGCTTACC 5495  
3399 CGATGTAACCACTCGTGAACCACTGATCTTACGATCTTACGATCTTACGATCTTACG 3458  
5496 CGATGTAACCACTCGTGAACCACTGATCTTACGATCTTACGATCTTACGATCTTACG 5555  
3459 CTGCGTGAAGAAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3518  
5556 CTGCGTGAAGAAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 5615  
3519 AATGTTGAATATCT 3578  
5616 AATGTTGAATATCT 5675  
3579 GTCTCAAGAGCGGATCAATTTGAATTTGAATTTGAATTTGAATTTGAATTTGAATTTGA 5675  
5676 GTCTCAAGAGCGGATCAATTTGAATTTGAATTTGAATTTGAATTTGAATTTGAATTTGA 3638  
3639 GCACATTTCCCGAAAAAGTGCAC 3662  
5736 GCACATTTCCCGAAAAAGTGCAC 5759

RESULT 15

US-09-573-322-21  
Sequence 21, Application US/09573322  
Patent No. 6531289  
GENERAL INFORMATION:  
APPLICANT: Bradley, John D.  
APPLICANT: Thompson, Craig M.  
APPLICANT: Moore, Jeffrey B.  
APPLICANT: Mobbe, C. Richard  
APPLICANT: Bailey, David A.  
TITLE OF INVENTION: Regulated Gene Expression in Yeast and  
FILE REFERENCE: Methods of Use  
CURRENT APPLICATION NUMBER: US/09/573,322  
PRIOR FILING DATE: 2000-05-18  
PRIOR APPLICATION NUMBER: 09/404,066  
PRIOR FILING DATE: 1999-09-23  
PRIOR APPLICATION NUMBER: 09/138,024  
PRIOR FILING DATE: 1998-08-21  
PRIOR APPLICATION NUMBER: 60/056,719  
NUMBER OF SEQ ID NOS: 30  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 21  
LENGTH: 7333  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: ZM197 plasmid  
US-09-573-322-21

Query Match 59.3%; Score 2172.8; DB 4; Length 7333;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 2177; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

1479 GAGTTACCCCGGCGCGGCTTGGCGTATCATGATCATGATGCTTCTGCTGATAATT 1538  
3576 GAGTGAACCGGACGACGAGCTTGGCGTATCATGATCATGATGCTTCTGCTGATAATT 3635  
1539 GTTATCCGCTCAACAATCCACACATACGAGCGGAGAACATTAAGTGTAAAGCTCGG 1598  
3636 GTTATCCGCTCAACAATCCACACATACGAGCGGAGAACATTAAGTGTAAAGCTCGG 3695  
1599 GTGCTTAATGATGAGCTAACTCAATTAATGCTGAGCTGACATGCGCTTCCAGT 1658  
3696 GTGCTTAATGATGAGCTAACTCAATTAATGCTGAGCTGACATGCGCTTCCAGT 3755  
1659 CCGGAAAACTGTGTGCTGCTGATGATTAATGATGATGATGATGATGATGATGATGATG 1718  
3756 CCGGAAAACTGTGTGCTGCTGATGATTAATGATGATGATGATGATGATGATGATGATG 3815  
1719 TGGTATTGGGCGCTCTTCCGCTTCTGCTCACTGATCTGCTGAGCTGCTGCTGCTGCTG 3875  
3816 TGGTATTGGGCGCTCTTCCGCTTCTGCTCACTGATCTGCTGAGCTGCTGCTGCTGCTG 3875  
1779 TGGGCGAGCGGCTATCAGCTCACTCAAAAGCGGTAATACGTTATCCAGAAATCAAGG 1838  
3876 TGGGCGAGCGGCTATCAGCTCACTCAAAAGCGGTAATACGTTATCCAGAAATCAAGG 3935  
1839 ATTAAGCGAGAAAGCAATGTAAGCAAAAGCGCAAAAGCGGAGAACCGTAAAAAG 1898  
3936 ATTAAGCGAGAAAGCAATGTAAGCAAAAGCGGTAATACGTTATCCAGAAATCAAGG 3995  
1899 CCGGCTTCTGCGCTTCTTCCATAGGCTCCGCCCCCTGACGATCACTCAAAATCGAC 1958  
3996 CCGGCTTCTGCGCTTCTTCCATAGGCTCCGCCCCCTGACGATCACTCAAAATCGAC 4055  
1959 GCTCAAGTCAAGGCTGCGAAACCGGACAGACTTAAGATTAACAGGCTTCCCGCTG 4115  
4056 GCTCAAGTCAAGGCTGCGAAACCGGACAGACTTAAGATTAACAGGCTTCCCGCTG 2018  
2019 GAAGCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2078  
4116 GAAGCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4175



QY 2079 TTCTCCCTTGGGGAAGCGTGGCGCTTTCTCAATAGCTCAAGCTGATGATCTCAATGGG 2138  
 Db 4176 TTCTCCCTTGGGGAAGCGTGGCGCTTTCTCAATAGCTCAAGCTGATGATCTCAATGGG 4235  
 QY 2139 TGTAGTGTGTCGCTCAAGCTGGGCTGTGTGACGAACCCCGCTTACGCCGACCGCT 2198  
 Db 4236 TGTAGTGTGTCGCTCAAGCTGGGCTGTGTGACGAACCCCGCTTACGCCGACCGCT 4295  
 QY 2199 GCGCTTATCCGGTACTATGTCCTTGAATCCACCCGGTAAGACACATTTATGCGAC 2258  
 Db 4296 GCGCTTATCCGGTACTATGTCCTTGAATCCACCCGGTAAGACACATTTATGCGAC 4355  
 QY 2259 TGGGAGCAGCACTGGTAAACAGATTAACAGACGAGATTAAGGCGCTGTAACAGT 2318  
 Db 4356 TGGGAGCAGCACTGGTAAACAGATTAACAGACGAGATTAAGGCGCTGTAACAGT 4415  
 QY 2319 TCTTGAATGGTGGCCTTAACGCTACACTAGAGAGACAGTATTTGGTATCTGCGCTC 2378  
 Db 4416 TCTTGAATGGTGGCCTTAACGCTACACTAGAGAGACAGTATTTGGTATCTGCGCTC 4475  
 QY 2379 TGTGAAAGCCAGTTAAGCTTCCGAAAAAGAGTTGGTACTTTGATCCGGCAAAACCA 2438  
 Db 4476 TGTGAAAGCCAGTTAAGCTTCCGAAAAAGAGTTGGTACTTTGATCCGGCAAAACCA 4535  
 QY 2439 CCGCTGGTAGCGGTGGTCTTTTGTGTCAGACGACAGATTAACGGCGCAAAAAAGAT 2498  
 Db 4536 CCGCTGGTAGCGGTGGTCTTTTGTGTCAGACGACAGATTAACGGCGCAAAAAAGAT 4595  
 QY 2499 CTGAAGAAGATCTTTGATCTTTTCTACGCGGCTGACGCTCAGTGAACGAAAACTAC 2558  
 Db 4596 CTGAAGAAGATCTTTGATCTTTTCTACGCGGCTGACGCTCAGTGAACGAAAACTAC 4655  
 QY 2559 GTTAAAGGATTTGGTCAATGAGATTATCAAAAAAGATTTGACCTGATCTTTAAAT 2618  
 Db 4656 GTTAAAGGATTTGGTCAATGAGATTATCAAAAAAGATTTGACCTGATCTTTAAAT 4715  
 QY 2619 AAAAAAGATTTAAATCAATCTAAAGTATATAGATAAATTTGGCTGACAGTTACC 2678  
 Db 4716 AAAAAAGATTTAAATCAATCTAAAGTATATAGATAAATTTGGCTGACAGTTACC 4775  
 QY 2679 AATGCTTAATAGTAGAGCACTATCTCAGGATCTGTCTATTTGGTCAATCCATAGTTG 2738  
 Db 4776 AATGCTTAATAGTAGAGCACTATCTCAGGATCTGTCTATTTGGTCAATCCATAGTTG 4835  
 QY 2739 CTTGATCTCCCGTGTGTAGATACTACGATACGGAGGCTTACCATCTGCCCCAGTG 2798  
 Db 4836 CTTGATCTCCCGTGTGTAGATACTACGATACGGAGGCTTACCATCTGCCCCAGTG 4895  
 QY 2799 CTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACAGC 2858  
 Db 4896 CTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACAGC 4955  
 QY 2859 CAGCCGGAAGGGCCGAGCGCAGAAAGTGGTCTTGCACTTTATCCGCCCTCCATCCAGTCTA 2918  
 Db 4956 CAGCCGGAAGGGCCGAGCGCAGAAAGTGGTCTTGCACTTTATCCGCCCTCCATCCAGTCTA 5015  
 QY 2919 TTAATTTGTCGGGAAAGCTAGAGTAAGTATGTCGCAATTAAATAGTTTGGCAACGTTG 2978  
 Db 5016 TTAATTTGTCGGGAAAGCTAGAGTAAGTATGTCGCAATTAAATAGTTTGGCAACGTTG 5075  
 QY 2979 TTGCAATTTGCTACAGGCACTGCTGATGTCACGCTGCTGTTGGTATGCTTCAATTCAGCT 3038  
 Db 5076 TTGCAATTTGCTACAGGCACTGCTGATGTCACGCTGCTGTTGGTATGCTTCAATTCAGCT 5135  
 QY 3039 CCGGTTCCCAACGATCAAGCGAGATTAATGATCCCGCATTTGTGCAAAAAAGCGTTA 3098  
 Db 5136 CCGGTTCCCAACGATCAAGCGAGATTAATGATCCCGCATTTGTGCAAAAAAGCGTTA 5195  
 QY 3099 GCTCTTTCGGTCTTCCGATCGTTGTCAAGAGTAAGTTGGCCGAGTGTATCACTCATGG 3158  
 Db 5196 GCTCTTTCGGTCTTCCGATCGTTGTCAAGAGTAAGTTGGCCGAGTGTATCACTCATGG 5255

QY 3159 TTATGGAGCAGCTGATTAATCTCTCTACTGTCATGCCCATCCGTAAGATGCTTTCTGGA 3218  
 Db 5256 TTATGGAGCAGCTGATTAATCTCTCTACTGTCATGCCCATCCGTAAGATGCTTTCTGGA 5315  
 QY 3219 CTGGTGAATCTCAACCAAGTCAATTCAGAAATAGTATGATCGGCGACCGAGTTGCTCTT 3278  
 Db 5316 CTGGTGAATCTCAACCAAGTCAATTCAGAAATAGTATGATCGGCGACCGAGTTGCTCTT 5375  
 QY 3279 GCCCGGCGTCAATACGGGATTAATACCGGCGCACATAGCAAACTTTAAAGTCTCATCA 3338  
 Db 5376 GCCCGGCGTCAATACGGGATTAATACCGGCGCACATAGCAAACTTTAAAGTCTCATCA 5435  
 QY 3339 TTGAAAAACGTTCTTCGGGGGGAACCTCAAGATCTTACCGCTGTGAGATCCAGTT 3398  
 Db 5436 TTGAAAAACGTTCTTCGGGGGGAACCTCAAGATCTTACCGCTGTGAGATCCAGTT 5495  
 QY 3399 CGATGAACCACTCTGTCACCCCACTGATCTTCAAGCATCTTTTACTTTCAACGCGTTT 3458  
 Db 5496 CGATGAACCACTCTGTCACCCCACTGATCTTCAAGCATCTTTTACTTTCAACGCGTTT 5555  
 QY 3459 CTGGGTGACAAAAACAGAAAGCAAAATGCGCAAAAAAGGAAATTAAGGCGACACGGA 3518  
 Db 5556 CTGGGTGACAAAAACAGAAAGCAAAATGCGCAAAAAAGGAAATTAAGGCGACACGGA 5615  
 QY 3519 AATGTTGAATCTCATCTCTCTCTTTTCAATTAATGAAGATTTATACAGGTTAT 3578  
 Db 5616 AATGTTGAATCTCATCTCTCTCTTTTCAATTAATGAAGATTTATACAGGTTAT 5675  
 QY 3579 GTCTCATGAGCGGATTAATTTGAATGATTTTGAATAAATAAATAAGGCTTCCGC 3638  
 Db 5676 GTCTCATGAGCGGATTAATTTGAATGATTTTGAATAAATAAATAAGGCTTCCGC 5735  
 QY 3639 GCACATTTCCCGAAAAAGTGCAC 3662  
 Db 5736 GCACATTTCCCGAAAAAGTGCAC 5759

Search completed: January 26, 2005, 12:21:01  
 Job time : 362.259 secs

**This Page Blank (uspto)**



```

Db      3055 ATCCAGCTGATCAGATCATATGCTGGGCTCTTTTCCGGCTGATGATGCCCAAGC 3114
Qy      105 TGGGCTATCTGGGCATCGGGAGAGAACCCGCGCTTTTCCCGAGGTTAAACG 164
Db      3115 TGGGCTATCTGGGCATCGGGAGAGAACCCGCGCTTTTCCCGAGGTTAAACG 3174
Qy      165 GCATGGAAGAAGTTTGGCCGAGATGACTGCTGCTGATTAAGCTTGAACGAAACGACG 224
Db      3175 GCATGGAAGAAGTTTGGCCGAGATGACTGCTGCTGATTAAGCTTGAACGAAACGACG 3224
Qy      225 TTACCATGATGATTTGGGAAGGTGGCATGACGCTTTAACGTTAACTGTTGTT 284
Db      3235 TTACCATGATGATTTGGGAAGGTGGCATGACGCTTTAACGTTAACTGTTGTT 3294
Qy      285 CAGGCCACTGGGATATCAGATTGCTGCGGCTTTTCCGGAACAGTTCCGATGTCAGC 344
Db      3295 CAGGCCACTGGGATATCAGATTGCTGCGGCTTTTCCGGAACAGTTCCGATGTCAGC 3354
Qy      345 CCGAAGCGCATGACGACCCGAAACAAATACCGGCGACGCGGAACCTGCGCGGTG 404
Db      3355 CCGAAGCGCATGACGACCCGAAACAAATACCGGCGACGCGGAACCTGCGCGGTG 3414
Qy      405 CAGATTAAATGACAGCGGTGCGCGCTGGGATATTAGCTGACGAGACGCGATATCTGCGC 464
Db      3415 CAGATTAAATGACAGCGGTGCGCGCTGGGATATTAGCTGACGAGACGCGATATCTGCGC 3474
Qy      465 TGGATGCCGCGAAGAAATGACATGATATCCCGTGAATTAACCGGCGGCGCGCTGTC 524
Db      3475 TGGATGCCGCGAAGAAATGACATGATATCCCGTGAATTAACCGGCGGCGCGCTGTC 3534
Qy      525 ATTCAAGTTTTTGAACCCGCGGAGAGACGCGGAGACTGCGGATGTTAACTGTTAAACG 584
Db      3535 ATTCAAGTTTTTGAACCCGCGGAGAGACGCGGAGACTGCGGATGTTAACTGTTAAACG 3594
Qy      585 GTGATGAGCAGATGAAATGCTCGACAGCTGCGAAGACAGCAGCT 631
Db      3595 GTGATGAGCAGATGAAATGCTCGACAGCTGCGAAGACAGCAGCT 3641

```

```

RESULT 2
LOCUS   XXU02426                20125 bp    DNA    linear    SYN 29-JAN-1997
DEFINITION Cloning vector lambda EMBL3 SP6/T7, left arm.
ACCESSION U02426
VERSION   U02426.1 GI:413792
KEYWORDS SOURCE
ORGANISM Cloning vector lambda EMBL3 SP6/T7
REFERENCE Cloning vector lambda EMBL3 SP6/T7
AUTHORS   1 (bases 1 to 20125)
TITLE     Cloning vector lambda EMBL3 SP6/T7
JOURNAL   ClonTECH Vectors On Disc version 1.3
REFERENCE Unpublished
AUTHORS   2 (bases 1 to 20125)
TITLE     ClonTECH Vectors On Disc version 1.3
JOURNAL   Direct Submission
COMMENT   Submitted (07-OCT-1993) Paul A. Kites, ClonTECH Laboratories, Inc.,
          1020 East Meadow Circle, Palo Alto, CA 94303, USA
          This sequence has been compiled from information in the sequence
          databases, published literature and other sources. If you suspect
          there is an error in this sequence, please contact ClonTECH's
          Technical Service Department at (415) 424-8222 or (800) 662-2566,
          extension 3 or E-mail TECH@CLONTECH.COM.
FEATURES
source    1..20125
           /organism="Cloning vector lambda EMBL3 SP6/T7"
           /mol_type="genomic DNA"
           /db_xref="taxon:31784"
ORIGIN
Query Match      82.8%; Score 585.4; DB 12; Length 20125;
Best Local Similarity 99.8%; Pred. No. 2.3e-148;

```

```

Matches 586; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy      45 ATCTAGCTGATCAGATCATATGCTGGGCTCTTTTCCGGCTGATGATGCCCAAGC 104
Db      3055 ATCCAGCTGATCAGATCATATGCTGGGCTCTTTTCCGGCTGATGATGCCCAAGC 3114
Qy      105 TGGGCTATCTGGGCATCGGGAGAGAACCCGCGCTTTTCCCGAGGTTAAACG 164
Db      3115 TGGGCTATCTGGGCATCGGGAGAGAACCCGCGCTTTTCCCGAGGTTAAACG 3174
Qy      165 GCATGGAAGAAGTTTGGCCGAGATGACTGCTGCTGATTAAGCTTGAACGAAACGACG 224
Db      3175 GCATGGAAGAAGTTTGGCCGAGATGACTGCTGCTGATTAAGCTTGAACGAAACGACG 3224
Qy      225 TTACCATGATGATTTGGGAAGGTGGCATGACGCTTTAACGTTAACTGTTGTT 284
Db      3235 TTACCATGATGATTTGGGAAGGTGGCATGACGCTTTAACGTTAACTGTTGTT 3294
Qy      285 CAGGCCACTGGGATATCAGATTGCTGCGGCTTTTCCGGAACAGTTCCGATGTCAGC 344
Db      3295 CAGGCCACTGGGATATCAGATTGCTGCGGCTTTTCCGGAACAGTTCCGATGTCAGC 3354
Qy      345 CCGAAGCGCATGACGACCCGAAACAAATACCGGCGACGCGGAACCTGCGCGGTG 404
Db      3355 CCGAAGCGCATGACGACCCGAAACAAATACCGGCGACGCGGAACCTGCGCGGTG 3414
Qy      405 CAGATTAAATGACAGCGGTGCGCGCTGGGATATTAGCTGACGAGACGCGATATCTGCGC 464
Db      3415 CAGATTAAATGACAGCGGTGCGCGCTGGGATATTAGCTGACGAGACGCGATATCTGCGC 3474
Qy      465 TGGATGCCGCGAAGAAATGACATGATATCCCGTGAATTAACCGGCGGCGCGCTGTC 524
Db      3475 TGGATGCCGCGAAGAAATGACATGATATCCCGTGAATTAACCGGCGGCGCGCTGTC 3534
Qy      525 ATTCAAGTTTTTGAACCCGCGGAGAGACGCGGAGACTGCGGATGTTAACTGTTAAACG 584
Db      3535 ATTCAAGTTTTTGAACCCGCGGAGAGACGCGGAGACTGCGGATGTTAACTGTTAAACG 3594
Qy      585 GTGATGAGCAGATGAAATGCTCGACAGCTGCGAAGACAGCAGCT 631
Db      3595 GTGATGAGCAGATGAAATGCTCGACAGCTGCGAAGACAGCAGCT 3641

```

```

RESULT 3
LOCUS   CVU39284                42529 bp    DNA    linear    SYN 13-APR-1996
DEFINITION Cloning vector TlP97-1, lambda phage lacZ translational fusion
ACCESSION U39284
VERSION   U39284.1 GI:1066304
KEYWORDS SOURCE
ORGANISM Cloning vector TlP97-1
REFERENCE Cloning vector TlP97-1
AUTHORS   1 (bases 1 to 42529)
TITLE     St Pierre, R. and Linn, T.
          A refined vector system for the in vitro construction of
          single-copy transcriptional or translational fusions to lacZ
          Gene 169 (1), 65-68 (1996)
JOURNAL   MEDLINE
PUBMED    8635751
REFERENCE 2 (bases 1 to 42529)
AUTHORS   St Pierre, R.
TITLE     Direct Submission
COMMENT   Submitted (24-OCT-1995) Thomas Linn, Microbiology and Immunology,
          University of Western Ontario, London, Ontario N6A5C1, Canada
FEATURES
source    1..42529
           /organism="Cloning vector TlP97-1"
           /mol_type="genomic DNA"
           /db_xref="taxon:43840"
           /note="rmbt2 terminator"
terminator

```

```
gene      complement(19507. .20760)
          /gene="lacZ"
          complement(19507. .20760)
          /gene="lacZ"
          /codon_start=1
          /product="lactose permease"
          /protein_id="AAC53648.1"
          /db_xref="GI:1066305"
          /translation="MYLKTNTFMWFGLLFFFFVFIFMGAVPPFPPIWLDINHISKSD
          TGLIFPAISLPSLFOPLFGLSDKLGIKRLIMITGMLWMPAPFIFIGPLQYN
          ILVGSIVGSIYLGCFNAGAPAVEAFIEKVRNSNFERGARMGCGMAGCAITVGI
          MFTINQFQWLGSCALLAVLFPKATDPSATYANANAGAHNSAFSLKALERL
          QPKWFLSYIGISCTYDFDQOPANFTSPFATGEGTVEFVITMGELNASTIN
          FPAFLINRIGKNAIILAGTISVRIIGSSFAISLEAVLILKTLHMEVDFLLVGC
          KYITQSEVERSATIYLVCFKFKOLAMFMSVLGNNYESIGFOGALVIVGLALGP
          TLISVFLSGPPLSLRQYNEVA"
          complement(20812. .23859)
          /gene="lacZ"
          complement(20812. .23859)
          /gene="lacZ"
          /note="EcoRI site normally present at the 3' end of lacZ
          has been removed by an uncharacterized mutation"
          /codon_start=1
          /translation="beta-galactosidase"
          /protein_id="AAC53649.1"
          /db_xref="GI:1066306"
          /translation="VILQRDMENGVLTQNLRLAHPFASMRNSEARTDPSQOQR
          SLNEMRFAMPAPAVESEWECDELPEADTVVPSNMOMHGDAPITNTVITYTN
          PEPVETNPQCYSLTFNVDESMLOEGOTRIIPGVNSAFHLMCGHGWGQDSRLP
          SEPLSAFIRAGENRLAVMTLRMSDGLSLENDQMRSGITRDVSLAKHTPTQSDPH
          VATRFNDFSRVALBAEVOMCGELRDYLRVSLMQGTQVASTAPRGGITIDERG
          YADVLTALINENPLMSASIPNLVRAVELTADGTLIEAEADVGEREYRILGGL
          LNKPEPLIRGVNHEHPLHQVMDQETQDILLKONFNVRSSHYNHMLWT
          LCDRYGVVDEANIETHGVNPNRLTDDPMILPMSRVPVSRQARBNHPSVILWL
          GNESGHNADNLRWIKSVDPSPVOEGGADTTATDIIICPMYARVDEQPPAP
          KMSIKKLSLPGETRPILCEYAHAMNSLQGRKAYQAFQYRILQGGFPMWDOS
          LIRKDENGPWSAIGDGDTPNDROFCMNGLVADRPALTEAHQOQFQRLS
          GQITLEVSEYLFRHSNDELHMVALLDQKPLAAGEVLPDVAPOKQULIELELPQPS
          AGOLMLTVRVVQPATAMSEAGHISAMQOMLAENLSVTLPAASHAIPHLTSSMDPS
          IELGNKQFNRQSGFLSOWMIGDKOLITPLRQOFTRAPLDNDIGVSEARLIDPMW
          VERKAGHQAQEAALIOCTADTLADAVLITTAHMOHGTFLFSRKYVLDSSGM
          AITVDVASTDTPHARIGLNCQLAQVAERVMNIGLGEQENTPRILTACDRNDL
          SDMTVPVFPSENLRCGTRBLNTPHQRWDIPFNISRYSQOQLMETSHRLLHABE
          GTWLNDGFHMGIGDSDSWSPSAEPQLSAGRYHYQLWVCOK"
          complement(23874. .23957)
          /note="multiple cloning site"
          complement(23973. .24141)
          /note="trnBT1 terminator"
          24475. .25665
          /gene="tecr"
          24475. .25665
          /gene="tecr"
          /codon_start=1
          /translation="tetraacycline resistance protein"
          /protein_id="AAC53650.1"
          /db_xref="GI:1066307"
          /translation="MKSNMALIVITGTVTLDAVIGLVMPLPGLRLDIYHSDSIASH
          YGVLLAYALMOPICAPVIGLSDRGRGRMPRLARPPVAGAGLISLH
          RIVAGITGATGAVAGATYADITDGEDRRHRLGLSAGCGVAGVAPAGGLAISLH
          APFLAAVLNGLNLILGCFLMQESHKGERMPRLARPPVAGAGLISLH
          FFIWQVGPVAPALMVI FGEDRFRMSATMIGLSAVGILHALLAQAVTGATRGF
          KOAIIAGMAADALGVLLAFATRGMAAFIPIILASGGLGIPALQAMISROVDDHGS
          QLOGSLAALTSITGPIVITAYASASTMNGLAMIVGAALVLVCLPLRRAMSR
          ATSR"
          33602. .35450
          /note="imm21 substitution"

mlec_feature
terminator
gene
CDS
```

```
ORIGIN
Query Match      82.8%; Score 585.4; DB 12; Length 42529;
Best Local Similarity 99.8%; Pred. No. 2.5e-148;
```

```
Matches 586; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 45 ATCTAGCTGCATCGATCATATGTCGGGCTTTTTCGGGCTCACTATGCCCAAGC 104
    |||
DB 3055 ATCCAGCTGCATCGATCATATGTCGGGCTTTTTCGGGCTCACTATGCCCAAGC 3114
QY 105 TGGCGCATATCGGGCATCGGGGAGAAAGAACCCGTCCTTTTCCCGGAGGTTGAAGCG 164
    |||
DB 3115 TGGCGCATATCGGGCATCGGGGAGAAAGAACCCGTCCTTTTCCCGGAGGTTGAAGCG 3174
QY 165 GCATGAAAGAGTTTGGCGAGAGTACTGCTGCTGATGACCTTGAGCGGAAAGCAGC 224
    |||
DB 3175 GCATGAAAGAGTTTGGCGAGAGTACTGCTGCTGATGACCTTGAGCGGAAAGCAGC 3224
QY 225 TTACATGATGATTCGGGAAAGGTGCGCAATGACGCCCTTTAAAGGTGAACCTGTCGT 284
    |||
DB 3235 TTACATGATGATTCGGGAAAGGTGCGCAATGACGCCCTTTAAAGGTGAACCTGTCGT 3294
QY 285 CAGGCCACCTGGGATACAGTTCGTCGGGCTTTTCCGGAACAGCTTCCGGAATGTCACG 344
    |||
DB 3295 CAGGCCACCTGGGATACAGTTCGTCGGGCTTTTCCGGAACAGCTTCCGGAATGTCACG 3354
QY 345 CCGAAGCGCATCAGCAACCCGAAACATACCGGCGACAGCCGGAACCTGCGGTGTG 404
    |||
DB 3355 CCGAAGCGCATCAGCAACCCGAAACATACCGGCGACAGCCGGAACCTGCGGTGTG 3414
QY 405 CAGATTATGACAGCGGTGCGGCGCTGGATATTACGTACAGCGAGACCGGATCTTGGC 464
    |||
DB 3415 CAGATTATGACAGCGGTGCGGCGCTGGATATTACGTACAGCGAGACCGGATCTTGGC 3474
QY 465 TGGATGCGCGGAAATGGAATGATATCCCGTAGATTACCCCGGCGGCGCTCGTTC 524
    |||
DB 3475 TGGATGCGCGGAAATGGAATGATATCCCGTAGATTACCCCGGCGGCGCTCGTTC 3534
QY 525 ATTACCGTTTTCGAACCCGCGAGAGAGCGGCGAGCTCGCGGTGCAATGTGTTTACAGC 584
    |||
DB 3535 ATTACCGTTTTCGAACCCGCGAGAGAGCGGCGAGCTCGCGGTGCAATGTGTTTACAGC 3594
QY 585 GTGATGAGCAGATGAAGATGCTTCGACAGCGCTGCAAGCAGCAGCT 631
    |||
DB 3595 GTGATGAGCAGATGAAGATGCTTCGACAGCGCTGCAAGCAGCAGCT 3641

RESULT 4
CVU39285 42530 bp DNA linear SYN 13-APR-1996
LOCUS Cloning vector TLp97-2, phage lambda lacZ translational fusion
DEFINITION vector, complete sequence.
ACCESSION U39285
VERSION U39285.1 GI:1066308
KEYWORDS Cloning vector TLp97-2
SOURCE artificial sequences; vector.
ORGANISM
REFERENCE 1 (bases 1 to 42530)
AUTHORS St Pierre,R. and Linn,T.
TITLE A refined vector system for the in vitro construction of
JOURNAL single-copy transcriptional or translational fusions to lacZ
MEDLINE Gene 169 (1), 65-68 (1996)
PUBMED 8635751
REFERENCE 2 (bases 1 to 42530)
AUTHORS ScPiere,R.
TITLE Direct Submision
JOURNAL Submitted (24-OCT-1995) Thomas Linn, Microbiology and Immunology,
University of Western Ontario, London, Ontario N6ASCI, Canada
FEATURES
SOURCE
1. .42530
/organism="Cloning vector TLp97-2"
/mol_type="genomic DNA"
/db_xref="taxon:43841"
complement(19404. .19489)
/note="trnBT2 terminator"
```





```
gene      complement(19507. .20760)
          /gene="lacc"
CDS        complement(19507. .20760)
          /gene="lacc"
          /codon_start=1
          /transl_table=1
          /product="lactose permease"
          /protein_id="AAC53654.1"
          /db_xref="GI:1066313"
          /translacion="MYLKNNTFMFGLFFFFYPFGAVPFPPIWLHNDINHSKD
          TGLIPALISLPSLFOPLFGLSDKLGIRKLLMIITGLVMPAPFPIFGPLLOYN
          ILVGIIVGIVGICFAGNAGAPAVAFIEKVRBSNFBFGAMGCGMAGCAIIVGI
          MFINOQFVFIAGSCALILAVLPFAKTDPSSATYANANGAHNSPSLKALIELR
          OPLKMLISLYIGVSCYDYVFDQOPAFPTSPFATGCGRVPRVYVTTMGLNASTM
          FPAPLIINRIGGKALILAGTISVRITIGSFATSALEVLTKLHMPEVFTLVGCP
          KFTYSGEVRFSATILVCFCEFRQLAMIEMSVLAGNMVESIGFQAGLVVLVAGLP
          TLISVFTLSGSPSLTRQYNEVA"
          complement(20812. .23859)
gene      /gene="lacc2"
          /note="EcoRI site normally present at the 3' end of lacc2
          has been removed by an uncharacterized mutation"
CDS        /codon_start=1
          /transl_table=1
          /product="beta-galactosidase"
          /protein_id="AAC53655.1"
          /db_xref="GI:1066314"
          /translacion="VVLQRDMENPVYTLNRLAHPFASWRSEARTRDPSQQLR
          SLNEMRFPAPFAVEAVESWLECDLPEADTVVVPNSMOMHGYDAPITYNTVYTIN
          PPVPTNPQCYSLITFNVDSEMLQEGOTRIIPGVNSAPHLMGNGMVGQDSRIP
          SERPLSFILAGNERLAVMTLRMSDGYLEQDMRRSGIFRVDYSLHKPTQISDPA
          VAFRNDPSRAVLEAVOMCGELDLTVLVSIMQETQVASTAIPFGGILIDERG
          YADVTLRLNENPKMSAEIPNLVRAVELHTADGTLIEAADCDFRREVRLEGL
          LNKRPILIRGVNHEHPLHQVMEQVODILLKONNFVARSHPNENHLYLT
          LCDRYGLVVDENALIEHGMVPMNRITDDPRLPAMSEVTRMVRQBRNSVITML
          GNESGHANDALYRMTKSVDPSPRYOEGGADTTATDIIICPYARVDEQPPRAP
          KWSIKKLSLPGETRPLILCYAHAMNSLGGFKTKQAFRQYRILDGGRFVMDVDS
          LIKTDENGPWSAYGDFGDPNDROFCMGLVFPADTRPHALTEAHHQOQFQFRLS
          GQITIEVSEYLFRHSDBELHMVYLDGKPLASGEVPLDVAPOKQULIELEFQPS
          AGOLMLTVRVVQPAATAMSEAGHISAMQOMLAENLVTLPASHAIPHLTISMDPC
          IELGNKMOFRQSGPLSOWMIGDKKOLLPLRPOPTRAPLDNDIGVSEATRIDPMM
          VERKAGHQAEALLOCTADTLADAVLITTAHMOGCTLISKRYRIDSSGOM
          AITVDEVASDTPHARIQLNQLQVAYERVNNGLDQENETPURLTACGRDNDPL
          SDMTVPVFPSENLRCGTRELNYPGQWRQDFQFNISRSQOQLMETSHRHLHAER
          GTMLNDGFHMGIGDSDSPSVAEPQLSAGRHYOLVWCOK"
          complement(23876. .23959)
misc_feature /note="multiple cloning site"
terminator complement(23975. .24143)
          /note="TribTI terminator"
          24477. .25667
gene      /gene="tccr"
          24477. .25667
CDS        /gene="tccr"
          /codon_start=1
          /transl_table=1
          /product="tetracycline resistance protein"
          /protein_id="AAC53656.1"
          /db_xref="GI:1066315"
          /translacion="MKSNALIVITGTVTLDAVIGLVMPVLPGLLDIVHSDSIASH
          YGVLLAYALMOPLCAPVGLASDRPGRRRPPLASLGATIDVIAVATVTVLVING
          RIVAGITGATGAVAGATYADITDGEDBDRRRPPLASLGATIDVIAVATVTVLVING
          APFLAAVAVLNLNLILGCFPLMQESHKGERMPRARNPVSPFMAAGMTIVALLMV
          FFIQVLGVGPALMLVIFGEBDRPMSATMGSLAVGGLIHALAQAVTGTATRGFB
          KOALIIAGMAADALGYVLLAFATRGMAFPMIILLASGGLIHALAQAVTGTATRGFB
          OLOGLSLAALTSITGPLYTAYASASATWGLAMIVGALVIVCLPLALRRAMSR
          ATSTP"
          33604. .35452
misc_feature /note="imm21 substructure"
```

Query Match 82.8%; Score 585.4; DB 12; Length 42531;  
Best Local Similarity 99.8%; Pred. No. 2.5e-148;

```
Matches 586; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 45 ATCTAGCTGCATCAGATCATATCTGCGGCTCTTTTCCGAGCTCACTGCGCCAAAC 104
Db 3055 ATCCAGCTGCATCAGATCATATCTGCGGCTCTTTTCCGAGCTCACTGCGCCAAAC 3114
QY 105 TGGCGCATCTGCGGCATCGGGGAGGAAGAGCCGTCCTTTTCCGAGGCTTGAAGCG 164
Db 3115 TGGCGCATCTGCGGCATCGGGGAGGAAGAGCCGTCCTTTTCCGAGGCTTGAAGCG 3174
QY 165 GCATGAAGAAGTGTGCGGAGATGACTGCTGCTGATTAAGCTTGAACGAAACGACG 224
Db 3175 GCATGAAGAAGTGTGCGGAGATGACTGCTGCTGATTAAGCTTGAACGAAACGACG 3234
QY 225 TTATCCATGATGATTCGGGAAGGTGTGCGCATGCAAGCTTTTAAAGGTGAACCTGTCGT 284
Db 3235 TTATCCATGATGATTCGGGAAGGTGTGCGCATGCAAGCTTTTAAAGGTGAACCTGTCGT 3294
QY 285 CAGGCCACCTGGGATACCACTTCGTCGCGGCTTTTCCGACACAGCTTCCGAGTGTGACG 344
Db 3295 CAGGCCACCTGGGATACCACTTCGTCGCGGCTTTTCCGACACAGCTTCCGAGTGTGACG 3354
QY 345 CCGAAGCGCATCAGCAACCGAACAATACCGGACGACCGGAACTGCGCGGTG 404
Db 3355 CCGAAGCGCATCAGCAACCGAACAATACCGGACGACCGGAACTGCGCGGTG 3414
QY 405 CAGATTATGACACAGCGGTGCGGCTGGAATATTACGTCAAGAGACGCGGTATCTGGC 464
Db 3415 CAGATTATGACACAGCGGTGCGGCTGGAATATTACGTCAAGAGACGCGGTATCTGGC 3474
QY 465 TGGATGCGCGAAGAAATGGAATGATATACCCCGGAGTTACCCGCGGCGCGCTGTC 524
Db 3475 TGGATGCGCGAAGAAATGGAATGATATACCCCGGAGTTACCCGCGGCGCGCTGTC 3534
QY 525 ATTACGCTTTTGAACCCGTGAGAGACGCGGACAGCTCGCGATGCAATGTGTTTACAGC 584
Db 3535 ATTACGCTTTTGAACCCGTGAGAGACGCGGACAGCTCGCGATGCAATGTGTTTACAGC 3594
QY 585 GTGATGAGACAGATGAAGATGCTTGACACGCTGCAAGACGCGAGCT 631
Db 3595 GTGATGAGACAGATGAAGATGCTTGACACGCTGCAAGACGCGAGCT 3641

RESULT 6
LOCUS CVU37692 42704 bp DNA linear SYN 13-APR-1996
DEFINITION Cloning vector lambda Txf97, lac2 transcriptional fusion vector,
complete sequence.
ACCESSION U37692
VERSION U37692.1 GI:1051181
KEYWORDS Cloning vector lambda Txf97
SOURCE Cloning vector lambda Txf97
ORGANISM artificial sequences; vector.
REFERENCE 1 (bases 1 to 42704)
AUTHORS St Pierre,R. and Linn,T.
TITLE A refined vector system for the in vitro construction of
single-copy transcriptional or translational fusions to lac2
JOURNAL Gene 169 (1), 65-68 (1996)
MEDLINE 96186904
PUBMED 8635751
REFERENCE 2 (bases 1 to 42704)
AUTHORS St Pierre,R. and Linn,T.
TITLE Direct Submmission
JOURNAL Submitted (29-SEP-1995) Thomas Linn, Microbiology and Immunology,
University of Western Ontario, London, Ontario N6A5C1, Canada
FEATURES
SOURCE
1. 42704
/organism="Cloning vector lambda Txf97"
/mol_type="genomic DNA"
/db_xref="taxon:43300"
1. 42704
/note="EcoRI site normally present at this site has been
```



JOURNAL MEDLINE PUBMED REFERENCE AUTHORS TITLE	71209066 4931680 2 (bases 45493 to 45963) Imada,M. and Teugilta,A. Amino acid sequence of lambda phage endolysin Nature New Biol. 233, 230-231 (1971)	JOURNAL MEDLINE PUBMED REFERENCE AUTHORS TITLE	959843 13 (bases 35578 to 35667) Humayun,Z., Jeffrey,A. and Prashne,M. Completed DNA sequences and organization of repressor-binding sites in the operators of phage lambda J. Mol. Biol. 112 (2), 265-277 (1977)
JOURNAL MEDLINE PUBMED REFERENCE AUTHORS TITLE	73215915 4515613 4 (bases 38597 to 38672) Dahlberg,J.E. and Blattner,F.R. In vitro transcription products of lambda DNA: Nucleotide sequences and regulatory sites (in) Fox,C.F. and Robinson,W.S. (Eds.): VIRUS RESEARCH. PROCEEDINGS OF 1973 ION-UCLA SYMPOSIUM: 533-544; Academic Press, New York (1973)	JOURNAL MEDLINE PUBMED REFERENCE AUTHORS TITLE	875019 14 (bases 38610 to 38732) Scherer,G., Hobom,G. and Kossel,H. DNA base sequence of the po promoter region of phage lambda Nature 265 (5590), 117-121 (1977)
JOURNAL MEDLINE PUBMED REFERENCE AUTHORS TITLE	75185528 1095210 6 (bases 35583 to 35600) Kleid,D.G., Agarwal,K.L. and Khorana,H.G. The nucleotide sequence in the promoter region of the gene N in bacteriophage lambda J. Biol. Chem. 250 (14), 5574-5582 (1975)	JOURNAL MEDLINE PUBMED REFERENCE AUTHORS TITLE	77100320 834253 15 (bases 38041 to 38241) Roberts,T.M., Shimatke,H., Brady,C. and Rosenberg,M. Sequence of Cro gene of bacteriophage lambda Nature 270 (5634), 274-275 (1977)
JOURNAL MEDLINE PUBMED REFERENCE AUTHORS TITLE	75185528 1095210 6 (bases 35583 to 35600) Kleid,D.G., Agarwal,K.L. and Khorana,H.G. The nucleotide sequence in the promoter region of the gene N in bacteriophage lambda J. Biol. Chem. 250 (14), 5574-5582 (1975)	JOURNAL MEDLINE PUBMED REFERENCE AUTHORS TITLE	78071724 593347 16 (bases 27616 to 28935) Davies,R.W., Schreier,P.H. and Buchel,D.E. Nucleotide sequence of the attachment site of coliphage lambda Nature 270 (5639), 757-760 (1977)
JOURNAL MEDLINE PUBMED REFERENCE AUTHORS TITLE	75189495 167018 7 (bases 35434 to 35618) Dahlberg,J.E. and Blattner,F.R. Sequence of the promoter-operator proximal region of the major leftward RNA of bacteriophage lambda Nucleic Acids Res. 2 (9), 1441-1458 (1975)	JOURNAL MEDLINE PUBMED REFERENCE AUTHORS TITLE	593399 17 (bases 37206 to 37263; 37914 to 37970) Humayun,Z. DNA sequence at the end of the ci gene in bacteriophage lambda Nucleic Acids Res. 4 (7), 2137-2143 (1977)
JOURNAL MEDLINE PUBMED REFERENCE AUTHORS TITLE	76031664 1178525 8 (bases 37945 to 38018) Maniatis,T., Jeffrey,A. and Kleid,D.G. Nucleotide sequence of the rightward operator of phage lambda Proc. Natl. Acad. Sci. U.S.A. 72 (3), 1184-1188 (1975)	JOURNAL MEDLINE PUBMED REFERENCE AUTHORS TITLE	78011659 909767 18 (bases 27617 to 27934) Landy,A. and Ross,W. Viral integration and excision: structure of the lambda att sites Science 197 (4309), 1147-1160 (1977)
JOURNAL MEDLINE PUBMED REFERENCE AUTHORS TITLE	75158212 1055375 9 (bases 44588 to 44773) Sklar,J., Yot,P. and Weisman,S.M. Determination of genes, restriction sites, and DNA sequences surrounding the 6S RNA template of bacteriophage lambda Proc. Natl. Acad. Sci. U.S.A. 72 (5), 1817-1821 (1975)	JOURNAL MEDLINE PUBMED REFERENCE AUTHORS TITLE	331474 19 (bases 39062 to 39170) Demiston-Thompson,K., Moore,D.D., Kruger,K.E., Furch,M.E. and Blattner,F.R. Physical structure of the replication origin of bacteriophage lambda Science 198 (4321), 1051-1056 (1977)
JOURNAL MEDLINE PUBMED REFERENCE AUTHORS TITLE	76217847 1098044 10 (bases 37905 to 37989) Walz,A., Pirocta,V. and Inelchen,K. Lambda repressor regulates the switch between PR and P <sub>RM</sub> promoters Nature 262 (5570), 665-669 (1976)	JOURNAL MEDLINE PUBMED REFERENCE AUTHORS TITLE	78054731 929187 20 (bases 44467 to 44807) Sklar,J.L. Structure and function of two regions of DNA controlling the synthesis of prokaryotic RNAs Thesis (1977)
JOURNAL MEDLINE PUBMED REFERENCE AUTHORS TITLE	76267718 958438 11 (bases 37946 to 38039) Smith,G.R., Eisen,H., Reichardt,L. and Hedgepeth,J. Deletions of lambda phage locating a p <sub>RM</sub> mutation within the rightward operator Proc. Natl. Acad. Sci. U.S.A. 73 (3), 712-716 (1976)	JOURNAL MEDLINE PUBMED REFERENCE AUTHORS TITLE	78234064 354508 22 (bases 13 to 72; 48391 to 48502) Nichols,B.P. and Donelson,J.E. 178-Nucleotide sequence surrounding the cos site of bacteriophage lambda DNA J. Virol. 26 (2), 429-434 (1978)
JOURNAL MEDLINE PUBMED REFERENCE AUTHORS TITLE	76152323 1062780 12 (bases 35578 to 35667; 37903 to 38027) Prashne,M., Backman,K., Humayun,M.Z., Jeffrey,A., Maurer,R., Meyer,B. and Sauer,R.T. Autoregulation and function of a repressor in bacteriophage lambda Science 194 (4261), 156-161 (1976)	JOURNAL MEDLINE PUBMED REFERENCE AUTHORS TITLE	78197067 666898 23 (bases 37938 to 38016; 35589 to 35666) Flashman,S.M. Mutational analysis of the operators of bacteriophage lambda Mol. Gen. Genet. 166 (1), 61-73 (1978)
JOURNAL MEDLINE PUBMED REFERENCE AUTHORS TITLE	76271154 76271154 24 (bases 37990 to 38982) Schwarz,E., Scherer,G., Hobom,G. and Kossel,H. Nucleotide sequence of cro, cII and part of the O gene in phage	JOURNAL MEDLINE PUBMED REFERENCE AUTHORS TITLE	368570 24 (bases 37990 to 38982) Schwarz,E., Scherer,G., Hobom,G. and Kossel,H. Nucleotide sequence of cro, cII and part of the O gene in phage



DRPEPELNEAIPNDERDTPVAMATTLRLKLTJGBELLTLASROOLIDMMEADKYAGPL  
 LRSALPMGFADKSGAGERSRGIIAALGDPKRSRIVIVITTTGSAQATDERNRQIA  
 EIGASLIKHM"  
 5316..5735  
 rep\_origin  
 /note="ColE1"

Query Match 81.3%; Score 574.6; DB 12; Length 5969;  
 Best Local Similarity 99.3%; Pred. No. 1.7e-145;  
 Matches 577; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 45 ATCTAGCTGATCGATCATATCGTGGGCTTTTTCGGGCTCATGCGCCAGC 104  
 DB 3508 ATCCAGCTGCATCAGATCATATCGTGGGCTTTTTCGGGCTCATGCGCCAGC 3567  
 QY 105 TGGCGCTATCTGGGCTATCGGGAGAGAACCCGCTTTTTCGGGAGTTGAACG 164  
 DB 3568 TGGCGCTATCTGGGCTATCGGGAGAGAACCCGCTTTTTCGGGAGTTGAACG 3627  
 QY 165 GCATGGAAGAAGTTTGGCGAGATGACTGCTGCTGATGACGTTGAGCGAAACGACG 224  
 DB 3628 GCATGGAAGAAGTTTGGCGAGATGACTGCTGCTGATGACGTTGAGCGAAACGACG 3687  
 QY 225 TTACCATGATGATTCGGGAGAGTGTGGCCATGACAGCGCTTTAAACGGTGAAGTGTG 284  
 DB 3688 TTACCATGATGATTCGGGAGAGTGTGGCCATGACAGCGCTTTAAACGGTGAAGTGTG 3747  
 QY 285 CAGGCCACTGGGATACCAAGTTTGTGGGCTTTTTCGGGACAGATTCGGAGTGTGACG 344  
 DB 3748 CAGGCCACTGGGATACCAAGTTTGTGGGCTTTTTCGGGACAGATTCGGAGTGTGACG 3807  
 QY 345 CCGAAGGCGATACCAACCCGCAATACCGGCGACAGCGCGAACTGCGGCTGGCGGTG 404  
 DB 3808 CCGAAGGCGATACCAACCCGCAATACCGGCGACAGCGCGAACTGCGGCTGGCGGTG 3867  
 QY 405 CAGATTAATGACAGCGGTGCGCGCTGGGATATTACGTCAGCGAGAGCGGGTATCCTGAC 464  
 DB 3868 CAGATTAATGACAGCGGTGCGCGCTGGGATATTACGTCAGCGAGAGCGGGTATCCTGAC 3927  
 QY 465 TGGATGCCGCGAATAAGCATGTGATACCCGCTGATACCCGCGGCGCGCTCGTTG 524  
 DB 3928 TGGATGCCGCGAATAAGCATGTGATACCCGCTGATACCCGCGGCGCGCTCGTTG 3987  
 QY 525 ATTACGTTTTTGAACCCGTGAGAGAGCGGAGACTCGCGGTGCAATGTGTTTACAGC 584  
 DB 3988 ATTACGTTTTTGAACCCGTGAGAGAGCGGAGACTCGCGGTGCAATGTGTTTACAGC 4047  
 QY 585 GTGATGAGCAGATGAAGATGCTGACACGCTGACGAACAC 625  
 DB 4048 GTGATGAGCAGATGAAGATGCTGACACGCTGACGAACAC 4088

RESULT 9  
 AF434925 6182 bp DNA circular SYN 04-APR-2002  
 LOCUS  
 DEFINITION Expression vector 409-FOR, complete sequence.  
 ACCESSION AF434925  
 VERSION AF434925.1 GI:16904158  
 KEYWORDS  
 SOURCE Expression vector 409-FOR  
 ORGANISM  
 .  
 Expression vector 409-FOR  
 artificial sequences; vectors.  
 REFERENCE  
 1 (bases 1 to 6182)  
 Zieles, H. and Huynh, C.O.  
 Intron-dependent stimulation of marker gene expression in cultured  
 insect cells  
 Insect Mol. Biol. 11 (1), 87-95 (2002)  
 JOURNAL  
 MEDLINE  
 PUBMED  
 11841506  
 2 (bases 1 to 6182)  
 Zieles, H. and Huynh, C.O.  
 Direct Submission  
 JOURNAL  
 Submitted (17-OCT-2001) Medical Entomology Section, Laboratory of

Parasitic Diseases, National Institutes of Health, 4 Center Drive  
 MSC 0425, Bethesda, MD 20892-0425, USA  
 Location/Qualifiers  
 1..6182

## FEATURES

source  
 /organism="Expression vector 409-FOR"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:176967"  
 /note="insect-specific"  
 97..614  
 /note="Autographa californica nuclear polyhedrosis virus  
 h5 enhancer"  
 615..1201  
 /note="Autographa californica nuclear polyhedrosis virus  
 immediate early promoter"  
 1266..1398  
 /note="artificial intron"  
 1541..3193  
 /codon\_start=1  
 /product="firefly luciferase"  
 /protein\_id="AAJ30782.1"  
 /db\_xref="GI:16904159"

## CDS

promoter  
 intron  
 CDS  
 m1sc\_feature  
 /note="Heliothis virescens p63 chaperonin gene double  
 poly-adenylation sequence"  
 4381..5241  
 /product="beta-lactamase"  
 /protein\_id="AAJ30783.1"  
 /db\_xref="GI:16904160"  
 /translation="MSIQHPRVALIPFPAAFCULPVFAPBETLVKVDADQLGARVY  
 IEDLNSGKILSFREPERPMMSFTVLGCAVLSRIDGQELGRIRHYSQNDLVE  
 YSPVTEKGLTIDGTVRELCSAALTMSNTNLANLLITLGPKEKTLFLHMGHVTPL  
 DRPEPELNEAIPNDERDTPVAMATTLRLKLTJGBELLTLASROOLIDMMEADKYAGPL  
 LRSALPMGFADKSGAGERSRGIIAALGDPKRSRIVIVITTTGSAQATDERNRQIA  
 EIGASLIKHM"  
 5529..5948  
 /note="ColE1"

## CDS

Query Match 81.3%; Score 574.6; DB 12; Length 6182;  
 Best Local Similarity 99.3%; Pred. No. 1.7e-145;  
 Matches 577; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

## ORIGIN

QY 45 ATCTAGCTGATCGATCATATCGTGGGCTTTTTCGGGCTCATGCGCCAGC 104  
 DB 3721 ATCCAGCTGCATCAGATCATATCGTGGGCTTTTTCGGGCTCATGCGCCAGC 3780  
 QY 105 TGGCGCTATCTGGGCTATCGGGAGAGAACCCGCTTTTTCGGGAGTTGAACG 164  
 DB 3781 TGGCGCTATCTGGGCTATCGGGAGAGAACCCGCTTTTTCGGGAGTTGAACG 3840  
 QY 165 GCATGGAAGAAGTTTGGCGAGATGACTGCTGCTGATGACGTTGAGCGAAACGACG 224  
 DB 3841 GCATGGAAGAAGTTTGGCGAGATGACTGCTGCTGATGACGTTGAGCGAAACGACG 3900  
 QY 225 TTACCATGATGATTCGGGAGAGTGTGGCCATGACAGCGCTTTAAACGGTGAAGTGTG 284  
 DB 3901 TTACCATGATGATTCGGGAGAGTGTGGCCATGACAGCGCTTTAAACGGTGAAGTGTG 3960  
 QY 285 CAGGCCACTGGGATACCAAGTTTGTGGGCTTTTTCGGGACAGATTCGGAGTGTGACG 344  
 DB 3961 CAGGCCACTGGGATACCAAGTTTGTGGGCTTTTTCGGGACAGATTCGGAGTGTGACG 4020

QY 345 CCGAAGCGCATCAGCAACCCGAAATATCCGCGACAGCCGGAACCTGCCGTGCCGTG 404  
DB 4021 CCGAAGCGCATCAGCAACCCGAAATATCCGCGACAGCCGGAACCTGCCGTGCCGTG 4080  
QY 405 CAGATTATATGACACCGGTGCGCGCTGGATATTATGATGAGAGAGACGGTATCTTGGC 464  
DB 4081 CAGATTATATGACACCGGTGCGCGCTGGATATTATGATGAGAGAGACGGTATCTTGGC 4140  
QY 465 TGGATGCCGCAAGAAATGATGATGATACCCCGTGAAGTTACCCGCGCGCGCGCTGTTTC 524  
DB 4141 TGGATGCCGCAAGAAATGATGATGATACCCCGTGAAGTTACCCGCGCGCGCGCTGTTTC 4200  
QY 525 ATTACAGTTTATTAACCCGTGAGAGACGGGACAGCTCGCGGTGGAATATGTTTTCACGC 584  
DB 4201 ATTACAGTTTATTAACCCGTGAGAGACGGGACAGCTCGCGGTGGAATATGTTTTCACGC 584  
QY 585 GTGATGAGACAGATGAAGATGCTGACACCGCTGCAAGAAC 625  
DB 4261 GTGATGAGACAGATGAAGATGCTGACACCGCTGATGAAGAAC 4301

RESULT 10  
AF434926 6182 bp DNA circular SYN 04-Apr-2002  
LOCUS AF434926  
DEFINITION Expression vector 409-MUT, complete sequence.  
ACCESSION AF434926  
VERSION AF434926.1 GI:16904161  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
Zieler,H. and Huynh,C.O.  
Intron-dependent stimulation of marker gene expression in cultured insect cells  
Insect Mol. Biol. 11 (1), 87-95 (2002)  
JOURNAL  
MEDLINE  
PUBMED  
21830665  
2 (bases 1 to 6182)  
Zieler,H. and Huynh,C.O.  
Direct Submission  
Submitted (17-OCT-2001) Medical Entomology Section, Laboratory of Parasitic Diseases, National Institutes of Health, 4 Center Drive MSC 0425, Bethesda, MD 20892-0425, USA  
FEATURES  
source  
1.. 6182  
/organism="Expression vector 409-MUT"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:176968"  
/note="insect-specific"  
97.. 614  
/note="Autographa californica nuclear polyhedrosis virus h75 enhancer"  
615.. 1201  
/note="Autographa californica nuclear polyhedrosis virus immediate early promoter"  
1266.. 1398  
/note="artificial intron with mutated 3' splice site"  
1541.. 3193  
/codon\_start=1  
/product="firefly luciferase"  
/protein\_id="AAJ30784.1"  
/db\_xref="GI:16904162"  
/translation="MEDAKNIKKGPAPYPLEDDTGAAGOLHKKMKRYALVPGTIAFTD  
AHEVITVAEYFEKMSVRLAEKMYGANTHRIIVGSENSLQFPMVLAFLIGVAV  
APANDIYNERELNSMNI SOPVVFSGKKI OKILNVOKPLIIOKI IIMDSKTVOG  
FQSMYTFVSHLP PGFNEDYDPVPS PPDKDTIALINSSGSLPGVAVLPRITACR  
FSHARDPIFGNOIIPDTALISLVPFHNGFQFTLGLICGFVVMVLPFEELPFS  
LQDYKIQSALLVPTLFSFRASSTLIDRYDLSNHEIASGAPLVKSGVAVAKRPLP  
GIRQGYGLTETTSALITLPEBDDPRGAVGVVFPFPAKVVDDLTGKTCGNRGSLC  
RGPMMSGYVNNPFAITALIDKQMLHSGDIAVMDHEPFIYDTRKSLIKIKGTOVA  
PALESTILQHPNIPDAGVAGLPDDDAGSLPAVVVLHEGKTYDTRKSLIKIKGTOVA  
AKKLRGGVVFVDEVPKGLTGKIDARKIRELILIAKKGSKL"

misc\_feature 3359.. 3464  
/note="Heliothis virescens p63 chaparonin gene double poly-adenylation sequence"  
4381.. 5241  
/codon\_start=1  
/product="beta-lactamase"  
/protein\_id="AAJ30785.1"  
/db\_xref="GI:16904163"  
/translation="MSIQHFRVALIPEPAFLPVPAPFETLVKKAEDQZGARVGY  
YSFVTEKHLLDQMTVRELCSAAITMSDNTAANLILTTIGSEKREILFLHNGDVTLE  
DRNEPILNEAI PVDERTTTPVAMATTLTRKILTEILASRQULIMADKVAQPL  
LRSLAPAGWFIADKSGAGRGSRGIIAALGDPGRSRIVLYTTGSAITVDERRRQIA  
EIGASLIKHW"  
5529.. 55948  
/note="ColEI"

Query Match 81.3%; Score 574.6; DB 12; Length 6182;  
Best Local Similarity 99.3%; Pred. No. 1.7e-145;  
Matches 577; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
QY 45 ATCTAGTCATCATGATCATATGCTGGGCTCTTTTCCGCTCATGTCATGCCCAAC 104  
DB 3721 ATCTAGTCATCATGATCATATGCTGGGCTCTTTTCCGCTCATGTCATGCCCAAC 3780  
QY 105 TGGCGCTATCTGGCGATCGGAGAGAACCCGTCCTTTTCCGCGAGGTGAAGCG 164  
DB 3781 TGGCGCTATCTGGCGATCGGAGAGAACCCGTCCTTTTCCGCGAGGTGAAGCG 3840  
QY 165 GCATGAAAGATTGGCCGAGATGACTGCTGCTGATGATGAGGAGAAACGACAG 224  
DB 3841 GCATGAAAGATTGGCCGAGATGACTGCTGCTGATGATGAGGAGAAACGACAG 3900  
QY 225 TTACAGATGATGATGCGGAAGGTGTGGCATGACAGCCTTTAAGGTAAGTGTGTT 284  
DB 3901 TTACAGATGATGATGCGGAAGGTGTGGCATGACAGCCTTTAAGGTAAGTGTGTT 3960  
QY 285 CAGGCACCTGGGATACAGTTCGTGCGGCGCTTTTCCGACACAGTTCGGATGTCAGC 344  
DB 3961 CAGGCACCTGGGATACAGTTCGTGCGGCGCTTTTCCGACACAGTTCGGATGTCAGC 4020  
QY 345 CCGAAGCGCATCAGCAACCCGAAATATCCGCGACAGCCGGAACCTGCCGTGCCGTG 404  
DB 4021 CCGAAGCGCATCAGCAACCCGAAATATCCGCGACAGCCGGAACCTGCCGTGCCGTG 4080  
QY 405 CAGATTATATGACACCGGTGCGCGCTGGATATTATGATGAGAGACGGTATCTTGGC 464  
DB 4081 CAGATTATATGACACCGGTGCGCGCTGGATATTATGATGAGAGACGGTATCTTGGC 4140  
QY 465 TGGATGCCGCAAGAAATGATGATGATACCCCGTGAAGTTACCCGCGCGCGCGCTGTTTC 524  
DB 4141 TGGATGCCGCAAGAAATGATGATGATACCCCGTGAAGTTACCCGCGCGCGCGCTGTTTC 4200  
QY 525 ATTACAGTTTATTAACCCGTGAGAGACGGGACAGCTCGCGGTGGAATATGTTTTCACGC 584  
DB 4201 ATTACAGTTTATTAACCCGTGAGAGACGGGACAGCTCGCGGTGGAATATGTTTTCACGC 584  
QY 585 GTGATGAGACAGATGAAGATGCTGACACCGCTGCAAGAAC 625  
DB 4261 GTGATGAGACAGATGAAGATGCTGACACCGCTGATGAAGAAC 4301

RESULT 11  
AF434927 6182 bp DNA circular SYN 04-Apr-2002  
LOCUS AF434927  
DEFINITION Expression vector 410-FOR, complete sequence.  
ACCESSION AF434927  
VERSION AF434927.1 GI:16904164  
KEYWORDS  
SOURCE  
ORGANISM  
Expression vector 410-FOR  
artificial sequences; vectors.



REFERENCE 1 (bases 1 to 6182)  
AUTHORS Ziebler, H. and Huynh, C.Q.  
TITLE Intron-dependent stimulation of marker gene expression in cultured insect cells  
JOURNAL Insect Mol. Biol. 11 (1), 87-95 (2002)  
MEDLINE 21830665  
PUBMED 11841506  
REFERENCE 2 (bases 1 to 6182)  
AUTHORS Ziebler, H. and Huynh, C.Q.  
TITLE Direct Submission  
JOURNAL Submitted (17-OCT-2001) Medical Entomology Section, Laboratory of Parasitic Diseases, National Institutes of Health, 4 Center Drive MSC 0425, Bethesda, MD 20892-0425, USA  
FEATURES  
Source  
1..6182  
/organism="Expression vector 410-REV"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:176969"  
/note="insect-specific"  
enhancer  
97..614  
/note="Autographa californica nuclear polyhedrosis virus h75 enhancer"  
promoter  
615..1201  
/note="Autographa californica nuclear polyhedrosis virus immediate early promoter"  
CDS  
1328..2980  
/codon\_start=1  
/product="firefly luciferase"  
/protein\_id="AAJ30786.1"  
/db\_xref="GI:16904165"  
/translation="MEDAKNTKKGPAPYPLEDGTAGROHLKAMRYALVPGTAFD AHENITVAYEFEMSVRLAMKRYGANTHRYVCSENSLQFFMPYGLPFGVAV APANDINERELNSWISQPTVAVSKGQKILNVGKPIIOKITIMDSKTDVQ FQSHADTFVTSHPGFEFEDVPSPFMDKTIALIMSSGSGLPKVALPRTACVR FSHARDPFGNOIIPDPAILSVFPFHGFGFTTGLYICGRVVLNFRFEEELFIS LODYKIOGALVPTLSEFPASTLIDKDLNLHSLASGAPLSKEVGEAVAKGPHLP GIDRGYGLTETTSAILITPEGDGKRGANVPPRAKVVDLDGKTLGVANORBELCY RGRPMGTYVNPRTATNALIDKOCMLHSGDIAYNDEDEHPTVDRLSKLITKGYQVA PALESLILQHPNITFDAGVAGLPDDAGELPAAVVLEHSGTMEKETVDVAVSOVTT AKKURGVVFDVEVPEKLTGKLDARKIREILIKAKGKSKL"  
intron  
3196..3328  
/note="artificial intron"  
misc\_feature  
3359..3464  
/note="heliothis virescens p63 chaperonin gene double poly-adenylation sequence"  
CDS  
4381..5241  
/codon\_start=1  
/product="beta-lactamase"  
/protein\_id="AAJ30787.1"  
/db\_xref="GI:16904166"  
/translation="MSIQHFRVALIIPPAFLCLPVFAHPETLVKVKDAEDQLGARVGY IELDLSGKILSRPREFRPMWSTFKVLGAVLSRIDAQEQLGRVHISQNDLVR YSPVTEKILDTGKTVRELCSAAITMSDNTANLLLTIGKRELTAFLHNNGDVHTR LDRPELKEAIPLNDERTPMVAATTRKLLTGLSELTLASRQOLIDMEADKVAGPL LRSALPAGWPIADKSGAGERSRGRIALGPDGKRSRLVVIYTTGSAQTMDERRQQLA EIGASLIKHW"  
FEATURES  
Source  
5259..5948  
/note="ColE1"  
rep\_origin  
Query Match 81.3%; Score 574.6; DB 12; Length 6182;  
Best Local Similarity 99.3%; Pred. No. 1.7e-145;  
Matches 577; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Db 3841 GCATGGAAAGAGTTTCCGAGAGATGACTGCTGCTGATTAACGTTAGCGAAACGACG 3900  
Qy 225 TTACCATGATGATTCGGGAAGGTGTGGCCATGACCGCTTTAAACGTGTAACGTT 284  
Db 3901 TTACCATGATGATTCGGGAAGGTGTGGCCATGACCGCTTTAAACGTGTAACGTT 3960  
Qy 285 CAGGCCACTGGGATACCAATTGCTGCGGCTTTTCCGGACACAGTTCCGGATGTCAGC 344  
Db 3961 CAGGCCACTGGGATACCAATTGCTGCGGCTTTTCCGGACACAGTTCCGGATGTCAGC 4020  
Qy 345 CCGAAGCGCATCAGAACCCGACCAATTCGGCGACACCGGAACTGCGCGGTG 404  
Db 4021 CCGAAGCGCATCAGAACCCGACCAATTCGGCGACACCGGAACTGCGCGGTG 4080  
Qy 405 CAGATTAAATGACACGCGTGGCGCGCTGGATATTACGTACAGCAGACCGGATCTGGC 464  
Db 4081 CAGATTAAATGACACGCGTGGCGCGCTGGATATTACGTACAGCAGACCGGATCTGGC 4140  
Qy 465 TGAATGCCGCAAAATGACATGATACCCCTGATTAACCCGCGGCGCGCTGCTTC 524  
Db 4141 TGAATGCCGCAAAATGACATGATACCCCTGATTAACCCGCGGCGCGCTGCTTC 4200  
Qy 525 ATTACGTTTTTGAACCGGTGAGAGACGGCGAGACTGGCGGTGAATGTGTTTACAGC 584  
Db 4201 ATTACGTTTTTGAACCGGTGAGAGACGGCGAGACTGGCGGTGAATGTGTTTACAGC 4260  
Qy 585 GTGATGAGACAGATGAAGATGCTGACACGCTGAGAAC 625  
Db 4261 GTGATGAGACAGATGAAGATGCTGACACGCTGATGAAC 4301  
RESULT 12  
LOCUS AF434931 6182 bp DNA circular SYN 04-APR-2002  
DEFINITION Expression vector 409-REV, complete sequence.  
ACCESSION AF434931  
VERSION AF434931.1 GI:16904176  
KEYWORDS  
SOURCE  
Expression vector 409-REV  
Expression vector 409-REV  
artificial sequences; vectors.  
REFERENCE  
1 (bases 1 to 6182)  
AUTHORS Ziebler, H. and Huynh, C.Q.  
TITLE Intron-dependent stimulation of marker gene expression in cultured insect cells  
JOURNAL Insect Mol. Biol. 11 (1), 87-95 (2002)  
MEDLINE 21830665  
PUBMED 11841506  
REFERENCE 2 (bases 1 to 6182)  
AUTHORS Ziebler, H. and Huynh, C.Q.  
TITLE Direct Submission  
JOURNAL Submitted (17-OCT-2001) Medical Entomology Section, Laboratory of Parasitic Diseases, National Institutes of Health, 4 Center Drive MSC 0425, Bethesda, MD 20892-0425, USA  
FEATURES  
Source  
1..6182  
/organism="Expression vector 409-REV"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:176973"  
/note="insect-specific"  
enhancer  
97..614  
/note="Autographa californica nuclear polyhedrosis virus h75 enhancer"  
promoter  
615..1201  
/note="Autographa californica nuclear polyhedrosis virus immediate early promoter"  
intron  
complement(1236..1368)  
/note="artificial intron"  
CDS  
1541..3193  
/codon\_start=1  
/product="firefly luciferase"  
/protein\_id="AAJ30794.1"



[illegible]

Query Match	Best Local Similarity	99.3%	Pred. No. 1.8e-145	Matches 577	Conservative	0	Mismatches 4	Indels 0	Gaps 0
45	ATTGAGCTGCATCAGATCATATCGTGGGCTTTTCCGGCTCAGTCAATCGCCCAAGC	104							
4960	ATCCAGCTGCATCAGATCATATCGTGGGCTTTTCCGGCTCAGTCAATCGCCCAAGC	5019							
105	TGGCGCTATCTGGGCATTCGGGAGGAAAGAACCCGTCCTTTTCCCGAGGTTGAAGCG	164							
5020	TGGCGCTATCTGGGCATTCGGGAGGAAAGAACCCGTCCTTTTCCCGAGGTTGAAGCG	5079							
165	GCATGGAAGAAGTTTCCGAGAGATGACTGCTGCTGATGACGTTGAGCCGAAACGACAGC	224							
5080	GCATGGAAGAAGTTTCCGAGAGATGACTGCTGCTGATGACGTTGAGCGAAGACGACAGC	5139							
225	TTTACCATGATGATTCGGGAAAGTGATGAGCCATGACGCTTAAACGGTGAACGTGTGTT	284							
5140	TTTACCATGATGATTCGGGAAAGTGATGAGCCATGACGCTTAAACGGTGAACGTGTGTT	5199							
285	CAGGCACTGGGATACAGATTCCTTCGCGCTTTTCCGAGACACAGTTCCGGATGTCAGC	344							
5200	CAGGCACTGGGATACAGATTCCTTCGCGCTTTTCCGAGACACAGTTCCGGATGTCAGC	5259							
345	CCGAAGCGCATCAGCAACCCGACCAATACCGGCGACAGCCGGAATCTCCCGTCCGCTGTG	404							
5260	CCGAAGCGCATCAGCAACCCGACCAATACCGGCGACAGCCGGAATCTCCCGTCCGCTGTG	5319							
405	CAGATTATATGACAGCGGTGGGCGCTGGGATATTAACGACAGCGAGACGGGTATCTGGGC	464							
5320	CAGATTATATGACAGCGGTGGGCGCTGGGATATTAACGACAGCGAGACGGGTATCTGGGC	5379							
465	TGATGTCGCGAGAATGGAACATGATATCCCGTGAATTAACCGCGGCGCGCTCTGTTCC	524							
5380	TGATGTCGCGAGAATGGAACATGATATCCCGTGAATTAACCGCGGCGCGCTCTGTTCC	5439							
525	ATTACAGTTTGAACCCGTGAGAGACGGGACAGCTCGCGGTGCAATGTGTTTACAGC	584							
5440	ATTACAGTTTGAACCCGTGAGAGACGGGACAGCTCGCGGTGCAATGTGTTTACAGC	5499							

```

Best Local Similarity 99.3%; Pred. No. 1,8e-145;
Matches 577; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 45 ATCTAGCTGCATCAGATCATTCCTCGCGGCTCTTTTTCCTCGGCTCAGTCATTCGCCAAGC 104
Db 5173 ATCCAGCTGCATCAGATCATTCATTCGTGGGTTCTTTTTCCTCGGCTCAGTCATTCGCCAAGC 5232
QY 105 TGGCGCTATCTCGGCTATCGGGAGAGAAAGCCCGTGCCTTTTTCCTCGGAGGTTGAAGC 164
Db 5233 TGGCGCTATCTCGGCTATCGGGAGAGAAAGCCCGTGCCTTTTTCCTCGGAGGTTGAAGC 5292
QY 165 GCATGGAAGAGTTTCCGAGGATGACCTGCTGCTGCACTTGAAGCGAAACGCAAGC 224
Db 5293 GCATGGAAGAGTTTCCGAGGATGACCTGCTGCTGCACTTGAAGCGAAACGCAAGC 5352
QY 225 TTACCATGATGATTCCTCGGAAAGGTGTGCCATGCAAGCCCTTTAAAGGTAAGTGTTCCTT 284
Db 5353 TTACCATGATGATTCCTCGGAAAGGTGTGCCATGCAAGCCCTTTAAAGGTAAGTGTTCCTT 5412
QY 285 CAGGCCACTGTGGATACCAAGTTGCTGCGGCTTTTCCGGACACAGTTCCGGATGTCAGC 344
Db 5413 CAGGCCACTGTGGATACCAAGTTGCTGCGGCTTTTCCGGACACAGTTCCGGATGTCAGC 5472
QY 345 CCGAAGCCATCAGCAACCCGAACATTAACCGCGCAGCGCGAATCTGCGCGCGGTCAGC 404
Db 5473 CCGAAGCCATCAGCAACCCGAACATTAACCGCGCAGCGCGAATCTGCGCGCGGTCAGC 5532
QY 405 CAGATTATACAGCGGTCGCGCTGCGATATTACGTCAGCGAGACCGGTAATCTGCGC 464
Db 5533 CAGATTATACAGCGGTCGCGCTGCGATATTACGTCAGCGAGACCGGTAATCTGCGC 5592
QY 465 TGAATGCCGCAAGAAATGGAATGGAATACCCGTAAGTTTACCCGCGCGGCTGCTGTC 524
Db 5593 TGAATGCCGCAAGAAATGGAATGGAATACCCGTAAGTTTACCCGCGCGGCTGCTGTC 5652
QY 525 ATTACGTTTTTGAACCCGTGAGGACCGGACAGCTCGCGGTGCAAAATGTTTTTACAGC 584
Db 5653 ATTACGTTTTTGAACCCGTGAGGACCGGACAGCTCGCGGTGCAAAATGTTTTTACAGC 5712
QY 585 GTGATGAGCAAGATGAAGATGCTGCACACCGCTCAGAAAC 625
Db 5713 GTGATGAGCAAGATGAAGATGCTGCACACCGCTCAGAAAC 5753

Search completed: January 26, 2005, 08:24:53
Job time : 3017.51 sec

```

Search completed: January 26, 2005, 08:24:53  
Job time : 3017.51 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 26, 2005, 02:11:40 ; Search time 2221.98 Seconds  
(without alignments)  
11594.581 Million cell updates/sec

Title: US-10-826-523-40

Perfect score: 1 ggcaccatgcgcacattc.....aaatcgacatggatcc 707

Sequence: 1 ggcaccatgcgcacattc.....aaatcgacatggatcc 707

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Database : EST:\*

1: gb\_est1:\*

2: gb\_est2:\*

3: gb\_est3:\*

4: gb\_est4:\*

5: gb\_est5:\*

6: gb\_est6:\*

7: gb\_est7:\*

8: gb\_est8:\*

9: gb\_est9:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	585.4	82.8	682	7	CK781302 UI-M-G10-
2	585.4	82.8	703	7	CF537771 UI-M-G10-
3	585.4	82.8	711	7	CF743678 UI-M-G10-
4	585.4	82.8	853	9	CR065566 Forward s
5	584.4	82.7	703	5	B0154655 NF087C06I
6	584.4	82.7	719	6	CD350897 UI-M-G10-
7	583.8	82.6	751	9	CR087413 Forward s
8	580.4	82.1	716	9	CR131675 Forward s
9	580.4	82.0	786	9	CR108810 Forward s
10	565.4	80.0	691	6	CD350776 UI-M-G10-
11	564.4	79.8	762	9	CR035207 Reverse s
12	563.4	79.7	749	6	CD351273 UI-M-G10-
13	562.4	79.5	733	9	CR139473 Forward s
14	555.4	78.5	633	9	CR167678 Reverse s
15	554.4	78.4	791	6	CB520716 UI-M-G10-
16	550.4	77.9	630	9	CR013093 Reverse s
17	549.4	77.7	702	9	CR084987 Reverse s
18	545.4	77.1	611	4	BR423083 EST533749
19	543.4	76.9	716	7	CK781284 UI-M-G10-
20	540.8	76.5	716	7	CF851544 PSMA007XC
21	537.2	76.0	706	1	AV731514 AV731514
22	535.4	75.7	730	9	CR077673 Reverse s
23	525.4	74.3	635	9	BX982794 Forward s
24	521.8	73.8	536	9	CR026633 Reverse s

C 25	516	73.0	887	9	CR192748 Forward s
C 26	515	72.8	885	9	CR255010 Forward s
C 27	498.4	70.5	683	6	CD346896 UI-M-FY0-
C 28	493	69.7	495	4	BG457760 NF036E03P
C 29	490.4	69.4	786	6	CB520098 UI-M-G10-
C 30	482.6	68.3	795	6	CD351738 UI-M-G10-
C 31	477.4	67.5	760	7	CF538326 UI-M-G10-
C 32	474.4	67.1	620	9	CR034585 Forward s
C 33	474.4	67.1	620	9	CR086282 Reverse s
C 34	474.4	67.1	620	9	CR171547 Reverse s
C 35	471.8	66.7	620	9	CR078862 Forward s
C 36	467.4	66.1	752	9	CR088896 Forward s
C 37	459.6	65.0	805	6	CD352530 UI-M-G10-
C 38	457	64.6	676	5	B0154673 NF095C11I
C 39	455.4	64.4	483	9	CR068828 Reverse s
C 40	447.4	63.3	547	9	CR148906 Forward s
C 41	447.4	63.3	885	9	CR024214 Reverse s
C 42	444.4	62.9	599	9	CR079016 Forward s
C 43	443.4	62.7	702	9	CR038019 Forward s
C 44	440.8	62.3	538	9	CR029330 Forward s
C 45	437.6	61.9	901	6	CA328434 UI-M-FY0-

#### ALIGNMENTS

RESULT 1  
LOCUS CK781302 682 bp mRNA linear EST 23-FEB-2004  
DEFINITION UI-M-G10-clp-g-05-0-UI.r1 NIH-BMAP\_G10 Mus musculus cDNA clone  
IMAGE:30619060 5', mRNA sequence.  
ACCESSION CK781302  
VERSION CK781302.1 GI:42746980  
KEYWORDS EST.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
REFERENCE NIH-MGC http://mgc.nci.nih.gov/  
AUTHORS 1 (bases 1 to 682)  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Published (1999)  
COMMENT Contact: Robert Strauberg, Ph.D.  
Email: ggaube-remail.nih.gov  
Tissue Procurement: Dr. James Lin, University of Iowa  
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa  
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Distribution information can be found at  
http://genome.uiowa.edu/distribution/mouse1.html  
This clone was contributed by the Brain Molecular Anatomy Project (BMAP)  
Seq primer: PYX-5.  
Location/Qualifiers  
1..682  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6"  
/db\_xref="taxon:10090"  
/clone\_id="IMAGE:30619060"  
/issue\_type="whole brain"  
/dev\_stage="embryo 13.5,14.5,16.5,17.5dpc"  
/lab\_host="DH10B (T1 phage resistant)"  
/notes="Organ: Brain; Vector: PYX-Asc; Site 1: Ecor I; Site 2: Not I; The library was constructed according to Bonaldi, Lemon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with Ecor I adaptor, digested with NotI and then cloned directionally into PYX-Asc vector. The library tag

## ORIGIN

sequence located between the Not I site and the polyA tail is AGCGAGACG. This library was created for the University of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemlin Chin, Ph.D., program coordinator."

Query Match 82.8%; Score 585.4; DB 7; Length 682;  
Best Local Similarity 99.8%; Pred. No. 2.3e-164;  
Matches 586; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

45 ATCTAGCTGCATAGAGATCATATCGCGGCTCTTTTTCGGCTAGTCATCGCCCAAGC 104  
DB 611 ATTCAGCTGCATAGAGATCATATCGCGGCTCTTTTTCGGCTAGTCATCGCCCAAGC 552  
QY 105 TGGCGCTATCTGGGCAATCGGAGAGAGAAAGCCCTGCTTTTCCGGAGGTTGAAGCG 164  
DB 551 TGGCGCTATCTGGGCAATCGGAGAGAGAAAGCCCTGCTTTTCCGGAGGTTGAAGCG 492  
QY 165 GATGAGAAAGATTGCGGAGAGATGCTGCTGCAATGACCTTGAAGCAAAACGACG 224  
DB 491 GATGAGAAAGATTGCGGAGAGATGCTGCTGCAATGACCTTGAAGCAAAACGACG 432  
QY 225 TTACCATGATGATTCGGAGAGGTCGCCATGACGCTTTAAACGTAAGTGTGCTT 284  
DB 431 TTACCATGATGATTCGGAGAGGTCGCCATGACGCTTTAAACGTAAGTGTGCTT 372  
QY 285 CAGGCGACCTGGATACCAAGTTCGCGCTTTTCCGGACACAGTTCGGATGTCAGC 344  
DB 371 CAGGCGACCTGGATACCAAGTTCGCGCTTTTCCGGACACAGTTCGGATGTCAGC 312  
QY 345 CCGAAGCGCATCAGCAACCCGAAACATACCGGAGACGCGAACTGCGCGGTGTG 404  
DB 311 CCGAAGCGCATCAGCAACCCGAAACATACCGGAGACGCGAACTGCGCGGTGTG 252  
QY 405 CAGATTATGACACCGCTGCGGCTGGAGATTTACGTCAGGAGAGAGCGGTATCTGCG 464  
DB 251 CAGATTATGACACCGCTGCGGCTGGAGATTTACGTCAGGAGAGAGCGGTATCTGCG 192  
QY 465 TGGATCCCGAGAAATGACATGATACCCCGTGAATACCCGCGGCGCGCTGCTG 524  
DB 191 TGGATCCCGAGAAATGACATGATACCCCGTGAATACCCGCGGCGCGCTGCTG 132  
QY 525 ATTCAGCTTTTGAACCCGTCGAGAGAGCGGAGACCTGCGCGTGAATGTTTACAGC 584  
DB 131 ATTCAGCTTTTGAACCCGTCGAGAGAGCGGAGACCTGCGCGTGAATGTTTACAGC 72  
QY 585 GTGATGAGACAGATGATGATGCTGACACCGCTGAGAGACCGCACT 631  
DB 71 GTGATGAGACAGATGATGATGCTGACACCGCTGAGAGACCGCACT 25

RESULT 2  
CF537771 703 bp mRNA linear EST 12-SEP-2003  
LOCUS UT-M-GIO-c-h-m-08-0-UI.r1 NIH BMAP\_GIO Mus musculus cDNA clone  
DEFINITION IMAGE:30537799 5', mRNA sequence.  
ACCESSION CF537771  
VERSION CF537771.1 GI:34589753  
KEYWORDS EST.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE  
AUTHORS 1 (bases 1 to 703)  
TITLE NIH-MGC <http://mgc.nci.nih.gov/>.  
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)  
COMMENT Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgabbs-remail.nih.gov](mailto:cgabbs-remail.nih.gov)  
Tissue Procurement: Dr. Jim Lin, University of Iowa  
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa

FEATURES  
Source  
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Distribution information can be found at  
<http://genome.uiowa.edu/distribution/mousecl.html>  
This clone was contributed by the Brain Molecular Anatomy Project  
(BMAP)

Seq primer: PYX-5.

Location/Qualifiers  
1..703

/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6"  
/db\_xref="taxon:10090"  
/clone="IMAGE:30537799"  
/issue\_type="whole brain"  
/dev\_stage="embryo 13.5,14.5,16.5,17.5dpc"  
/lab\_host="PH10B (T1 phase resistant)"  
/clone\_lib="NIH BMAP G10"  
/note="Organ: Brain; Vector: PYX-Asc; Site 1: EcoR I;  
Site 2: Not I; The library was constructed according  
to Bonaldi, Lennon and Soares, Genome Research, 6:791-806,  
1996. Denatured RNA was size fractionated on a 1% agarose  
gel. First strand cDNA synthesis was primed with oligo-dT  
primer containing a Not I site. Double strand cDNA was  
size selected according to RNA size fraction. Ligated  
directionally into PYX-Asc vector. The library tag  
sequence located between the Not I site and the polyA tail  
is AGCGAGACG. This library was created for the University  
of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the  
Developing Mouse Nervous System', supported by National  
Institute of Mental Health (NIMH), Hemlin Chin, Ph.D.,  
program coordinator."

## ORIGIN

Query Match 82.8%; Score 585.4; DB 7; Length 703;  
Best Local Similarity 99.8%; Pred. No. 2.3e-164;  
Matches 586; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

45 ATCTAGCTGCATAGAGATCATATCGCGGCTCTTTTTCGGCTAGTCATCGCCCAAGC 104  
DB 643 ATTCAGCTGCATAGAGATCATATCGCGGCTCTTTTTCGGCTAGTCATCGCCCAAGC 584  
QY 105 TGGCGCTATCTGGGCAATCGGAGAGAGAAAGCCCTGCTTTTCCGGAGGTTGAAGCG 164  
DB 583 TGGCGCTATCTGGGCAATCGGAGAGAGAAAGCCCTGCTTTTCCGGAGGTTGAAGCG 524  
QY 165 GATGAGAAAGATTGCGGAGAGATGCTGCTGCAATGACCTTGAAGCAAAACGACG 224  
DB 523 GATGAGAAAGATTGCGGAGAGATGCTGCTGCAATGACCTTGAAGCAAAACGACG 464  
QY 225 TTACCATGATGATTCGGAGAGGTCGCCATGACGCTTTTAAACGTAAGTGTGCTT 284  
DB 463 TTACCATGATGATTCGGAGAGGTCGCCATGACGCTTTTAAACGTAAGTGTGCTT 404  
QY 285 CAGGCGACCTGGATACCAAGTTCGCGCTTTTCCGAGACAGTTCGAGTGTGTCAGC 344  
DB 403 CAGGCGACCTGGATACCAAGTTCGCGCTTTTCCGAGACAGTTCGAGTGTGTCAGC 344  
QY 345 CCGAAGCGCATCAGCAACCCGAAACATACCGGAGACGCGAACTGCGCGGTGTG 404  
DB 343 CCGAAGCGCATCAGCAACCCGAAACATACCGGAGACGCGAACTGCGCGGTGTG 284  
QY 405 CAGATTATGACACCGCTGCGGAGAGGTCGCCATGACGCTTTTAAACGTAAGTGTGCTT 464  
DB 283 CAGATTATGACACCGCTGCGGAGAGGTCGCCATGACGCTTTTAAACGTAAGTGTGCTT 224  
QY 465 TGGATCCCGAGAAATGACATGATGCTGACACCGCTGAGAGACCGGATATCTGCGC 524  
DB 223 TGGATCCCGAGAAATGACATGATGCTGACACCGCTGAGAGACCGGATATCTGCGC 164  
QY 525 ATTCAGCTTTTGAACCCGTCGAGAGAGCGGAGACCTGCGCGTGAATGTTTACAGC 584



```

Db      163 ATTCACGTTTTTGAACCCGTCGAGAGAGCGGACACTCGCGGTGCAATGTTTTACAGC 104
Qy      585 GTGATGAGCAGATGAGATGCTGTGACACGCTGCGAAGACGACGCT 631
Db      103 GTGATGAGCAGATGAGATGCTGTGACACGCTGCGAAGACGACGCT 57

RESULT 3
CF743678/c
LOCUS   CF743678
DEFINITION
UI-M-G10-clg-e-19-0-UI.r1 NIH BMAP_G10 Mus musculus cDNA clone
IMAGE:30616722 5', mRNA sequence.
CF743678
ACCESSION
CF743678.1 GI:37640017
VERSION
EST.
KEYWORDS
Mus musculus (house mouse)
SOURCE
Mus musculus
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 711)
AUTHORS
NIH-MGC http://mgc.ncl.nih.gov/.
TITLE
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
Unpublished (1999)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-remail.nih.gov
Tissue Procurement: Dr. James Lin, University of Iowa
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
DNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
http://genome.uiowa.edu/distribution/mouse1.html
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)
Seq primer: PYX-5.
FEATURES
Location/Qualifiers
1..711
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:30616722"
/tissue_type="whole brain"
/dev_stage="embryo 13.5, 14.5, 16.5, 17.5dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_1lb="NIH_BMAP_G10"
/note="Organ: Brain; Vector: PYX-Asc; Site: 1: EcoR I;
Site 2: Not I; The library was constructed according
to Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
size selected according to mRNA size fraction, ligated
with EcoR I adaptor, digested with NotI and then cloned
directionally into PYX-Asc vector. The library tag
sequence located between the Not I site and the polyA tail
is AGCGAGACAG. This library was created for the University
of Iowa Brain Anatomy Project (BMAP). 'Gene Discovery in the
Developing Mouse Nervous System', supported by National
Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
program coordinator."
ORIGIN
Query Match 82.8%; Score 585.4; DB 7; Length 711;
Best Local Similarity 99.8%; Pred. No. 2,3e-164;
Matches 586; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 45 ATTCAGTCGATGAGATGATATGTCGGGCTTTTTCGGGTCAGTCATCGCCAGC 104
Db 654 ATTCAGTCGATGAGATGATATGTCGGGCTTTTTCGGGTCAGTCATCGCCAGC 595
Qy 105 TGGCGCTATCTGGGCAATCGGGAGAGAACCCGTCCTTTTCCCGAGATTGAAGC 164
Db 594 TGGCGCTATCTGGGCAATCGGGAGAGAACCCGTCCTTTTCCCGAGATTGAAGC 535

```

```

Qy 165 GCATGAAAGATTTCGCCAGATGACTGCTGCTGCACTTGACGTTGACCGAAAACGACG 224
Db 534 GCATGAAAGATTTCGCCAGATGACTGCTGCTGCTGCACTTGACGTTGACCGAAAACGACG 475
Qy 225 TTTCAGATGATGATGCGGAGAGGTCGCGATGCGACGCTTTTAACTGGTGAATGTTGGTT 284
Db 474 TTTCAGATGATGATGCGGAGAGGTCGCGATGCGACGCTTTTAACTGGTGAATGTTGGTT 415
Qy 285 CAGGCCACTGGATATACAGTTGCTGCGCGCTTTTTCGACACAGTTCCGATGTCGACG 344
Db 414 CAGGCCACTGGATATACAGTTGCTGCGCGCTTTTTCGACACAGTTCCGATGTCGACG 355
Qy 345 CCGAAGCGCATACGCAACCCGAAACATACCGGCGACACGCGGAACTCCGTCGGTGTG 404
Db 354 CCGAAGCGCATACGCAACCCGAAACATACCGGCGACACGCGGAACTCCGTCGGTGTG 295
Qy 405 CAGATTATGACAGCGGTGCGGCGCTGGATATTAGCTACGCGAGACGCGGTATCTGGC 464
Db 294 CAGATTATGACAGCGGTGCGGCGCTGGATATTAGCTACGCGAGACGCGGTATCTGGC 235
Qy 465 TGGATGCCGAGAAATGAGATGATGATCCCGTGAATTAACCCGCGGCGGCTGCTTC 524
Db 234 TGGATGCCGAGAAATGAGATGATGATCCCGTGAATTAACCCGCGGCGGCTGCTTC 175
Qy 525 ATTACGTTTTTTGACCCCGTGAGAGACGCGGACACTCGCGTCAATGTTTTACAGC 584
Db 174 ATTACGTTTTTTGACCCCGTGAGAGACGCGGACACTCGCGTCAATGTTTTACAGC 115
Qy 585 GTGATGAGCAGATGAGATGCTGACACGCTGAGAACCGACGCT 631
Db 114 GTGATGAGCAGATGAGATGCTGACACGCTGAGAACCGACGCT 68

RESULT 4
CRO65566
LOCUS   CRO65566
DEFINITION
Forward strand read from insert in 3'HPT insertion targeting and
chromosome engineering clone MHP250k12, genomic survey sequence.
CRO65566
ACCESSION
CRO65566.1 GI:4979156
VERSION
GSS: genome survey sequence; MISC.
KEYWORDS
Mus musculus (house mouse)
SOURCE
Mus musculus
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 853)
AUTHORS
Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,L.,
Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y.,
Rogers,J. and Bradley,A.
TITLE
Direct Submision
JOURNAL
Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. http://www.sanger.ac.uk/MISC
FEATURES
Location/Qualifiers
1..853
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/clone="MHP250k12"
/clone_1lb="MHP"
ORIGIN
Query Match 82.8%; Score 585.4; DB 9; Length 853;
Best Local Similarity 99.8%; Pred. No. 2,4e-164;
Matches 586; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 45 ATTCAGTCGATGAGATGATATGTCGGGCTTTTTCGGGTCAGTCATCGCCAGC 104
Db 264 ATTCAGTCGATGAGATGATATGTCGGGCTTTTTCGGGTCAGTCATCGCCAGC 323
Qy 105 TGGCGCTATCTGGGCAATCGGGAGAGAACCCGTCCTTTTCCCGAGATTGAAGC 164
Db 324 TGGCGCTATCTGGGCAATCGGGAGAGAACCCGTCCTTTTCCCGAGATTGAAGC 383

```

## ORIGIN

Query Match	82.7%;	Score 584.4;	DB 5;	Length 703
Best Local Similarity	99.7%;	Pred. No. 4.6e-164;		
Matches 585; Conservative	0;	Mismatches 6		

[illegible]

	719 bp	linear	EST 09-JUL-2003
UI-M-G10-cg9-a-18-0-UI.r1 NIH BMAP-G10	Mus musculus	CDNA clone	
IMAGE:6853243 5', mRNA sequence.			
CD350897			
CD350897.1 GI:3142412			
EST.			
Mus musculus (house mouse)			
Mus musculus			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			

REFERENCE 1 (bases 1 to 719)  
 NIH-MGC <http://mgc.nci.nih.gov/>.  
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
 TITLE Unpublished (1999)  
 JOURNAL COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgabbs-romail.nih.gov](mailto:cgabbs-romail.nih.gov)  
 Tissue Procurement: Dr. Jim Lin, University of Iowa  
 CDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
 CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
 Clone Distribution: Distribution information can be found at  
<http://genome.uiowa.edu/distribution/mouse1.html>  
 This clone was contributed by the Brain Molecular Anatomy Project (BMAP)

Seq primer: PYX-5.

## FEATURES

## Source

Location/Qualifiers  
 1..719  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="C57BL/6"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:6853243"  
 /isue\_type="whole brain"  
 /dev\_stage="embryo 13.5, 14.5, 16.5, 17, 5dpc"  
 /lab\_host="MD10B (T1 phage resistant)"  
 /clone\_lib="NIH BMAP G10"  
 /note="Organ: Brain; Vector: PYX-Asc; Site\_1: EcoR I; Site\_2: Not I; The library was constructed according to Bonaldi, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into PYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is AGCAGACAG. This library was created for the University of Iowa Brain Anatomy Project (BMAP). 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

## ORIGIN

Query Match 82.7%; Score 584.4; DB 6; Length 719;  
 Best Local Similarity 99.7%; Pred. No. 4.6e-164;  
 Matches 585; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

45 ATCTAGCTGCATCAGATCATATCGTGGGCTTTTTCGGGCTCAGTCATCGCCCAAGC 104  
 Db ATCTAGCTGCATCAGATCATATCGTGGGCTTTTTCGGGCTCAGTCATCGCCCAAGC 104  
 663 ATCCAGCTGCATCAGATCATATCGTGGGCTTTTTCGGGCTCAGTCATCGCCCAAGC 604  
 QY 105 TGGGCTATCTGGGCATCGGGAGAGAAAGCCGTCCTTTTCCCGAGGTTGAAGCG 164  
 Db 603 TGGGCTATCTGGGCATCGGGAGAGAAAGCCGTCCTTTTCCCGAGGTTGAAGCG 544  
 QY 165 GCATGGAAGAAGTTTGGCGAGATGACTGCTGCTGCACTTGAAGCGAAACGACG 224  
 Db 543 GCATGGAAGAAGTTTGGCGAGATGACTGCTGCTGCACTTGAAGCGAAACGACG 484  
 QY 225 TTTCACATGATGATCGGGAAAGTGTGGCCATGACGCTTTTAAAGGTAAGTGTGTT 284  
 Db 483 TTTCACATGATGATCGGGAAAGTGTGGCCATGACGCTTTTAAAGGTAAGTGTGTT 424  
 QY 285 CAGGCCACTGGGATACCAAGTTCGTGCGGCTTTTCCGACACAGTTCCGAGTGTGACG 344  
 Db 423 CAGGCCACTGGGATACCAAGTTCGTGCGGCTTTTCCGACACAGTTCCGAGTGTGACG 364  
 QY 345 CCGAAGGCAATCAGCAACCCGAAACAATACCGGCGACAGCGGAACTGCGGTGTG 404  
 Db 363 CCGAAGGCAATCAGCAACCCGAAACAATACCGGCGACAGCGGAACTGCGGTGTG 304  
 QY 405 CAGATTATGACAGCGGTGCGGCGCTGGGATATTACGTCAGCGGATCTCTGGC 464

Db 303 CAGATTATGACAGCGGTGCGGCGCTGGGATATTACGTCAGCGGATCTCTGGC 244  
 QY 465 TGAATGCCCGCAAAATGACATGTATACCCCTGAGTTACCCGCGGCGGCTCGTTG 524  
 Db 243 TGAATGCCCGCAAAATGACATGTATACCCCTGAGTTACCCGCGGCGGCTCGTTG 184  
 QY 525 ATTACGTTTTTGAACCCGTGAGAGAGGCGGAGCTGCGGCTGGAATGTGTTTACAGC 584  
 Db 183 ATTACGTTTTTGAACCCGTGAGAGAGGCGGAGCTGCGGCTGGAATGTGTTTACAGC 124  
 QY 585 GTGATGAGCAGATGATGATGCTGACACGCTGCGAGAACGACAGCT 631  
 Db 123 GTGATGAGCAGATGATGATGCTGACACGCTGCGAGAACGACAGCT 77

RESULT 7  
 CR087413/c 751 bp DNA linear GSS 05-JUL-2004  
 LOCUS Forward strand read from insert in 3'HPRT insertion targeting and  
 DEFINITION chromosome engineering clone MHP438f24, genomic survey sequence.  
 ACCESSION CR087413  
 VERSION GI:49821005  
 KEYWORDS GSS; genome survey sequence; MICE.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus

REFERENCE Bukaryota, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathia; Muridae; Murinae; Mus.  
 1 (bases 1 to 751)  
 AUTHORS Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,L.,  
 Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y.,  
 Rogers,J. and Bradley,A.

TITLE Direct Submission  
 JOURNAL Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,  
 CB10 1SA, UK. <http://www.sanger.ac.uk/MICE>

## FEATURES

## source

Location/Qualifiers  
 1..751  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:10090"  
 /clone="MHP438f24"  
 /clone\_lib="MHP"

## ORIGIN

Query Match 82.6%; Score 583.8; DB 9; Length 751;  
 Best Local Similarity 99.7%; Pred. No. 7e-164;  
 Matches 585; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

45 ATCTAGCTGCATCAGATCATATCGTGGGCTTTTTCGGGCTCAGTCATCGCCCAAGC 104  
 Db 623 ATCCAGCTGCATCAGATCATATCGTGGGCTTTTTCGGGCTCAGTCATCGCCCAAGC 564  
 QY 105 TGGGCTATCTGGGCATCGGGAGAGAAAGCCGTCCTTTTCCCGAGGTTGAAGCG 164  
 Db 563 TGGGCTATCTGGGCATCGGGAGAGAAAGCCGTCCTTTTCCCGAGGTTGAAGCG 504  
 QY 165 GCATGGAAGAAGTTTGGCGAGATGACTGCTGCTGCACTTGAAGCGAAACGACG 224  
 Db 503 GCATGGAAGAAGTTTGGCGAGATGACTGCTGCTGCACTTGAAGCGAAACGACG 444  
 QY 225 TTTCACATGATGATCGGGAAAGTGTGGCCATGACGCTTTTAAAGGTAAGTGTGTT 284  
 Db 443 TTTCACATGATGATCGGGAAAGTGTGGCCATGACGCTTTTAAAGGTAAGTGTGTT 384  
 QY 285 CAGGCCACTGGGATACCAAGTTCGTGCGGCTTTTCCGACACAGTTCCGAGTGTGACG 344  
 Db 383 CAGGCCACTGGGATACCAAGTTCGTGCGGCTTTTCCGACACAGTTCCGAGTGTGACG 324  
 QY 345 CCGAAGGCAATCAGCAACCCGAAACAATACCGGCGACAGCGGAACTGCGGTGTG 404  
 Db 323 CCGAAGGCAATCAGCAACCCGAAACAATACCGGCGACAGCGGAACTGCGGTGTG 264  
 QY 405 CAGATTATGACAGCGGTGCGGCGCTGGGATATTACGTCAGCGGATCTCTGGC 464

Db 263 CAGATTAAATGACACGGTGGCGCTGGGATATTACGTACAGGAGAGGGTATCCTGGC 204

QY 465 TGGATGCCGAGAAATGACATGATATACCCCGTGAATTACCCGGGGGGCGCGCTCTGTC 204

Db 203 TGGATGCCGAGAAATGACATGATATACCCCGTGAATTACCCGGGGGGCGCGCTCTGTC 524

QY 525 ATTACGTTTTTGAACCCGTGAGAGAGCGGAGACTCGCGGTGCAATGTGTTTTACAGC 144

Db 143 ATTACGTTTTTGAACCCGTGAGAGAGCGGAGACTCGCGGTGCAATGTGTTTTACAGC 584

QY 585 GTGATGAGCAGATGAAGATGCTCGACACGCTGCAAGACAGGAGCT 631

Db 83 GTGATGAGCAGATGAAGATGCTCGACACGCTGCAAGACAGGAGCT 37

RESULT 8  
CRI31675/c  
LOCUS  
DEFINITION Forward strand read from insert in 3'HPRT insertion targeting and  
ACCESSION CRI31675.1 GI:49879128  
VERSION CRI31675.1 GI:49879128  
KEYWORDS GSS; genome survey sequence; MICR.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
REFERENCE Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
AUTHORS 1 (bases 1 to 716)  
Adam, D.J., Biggs, P.J., Cox, A.V., Davies, R.M., van der Weyden, L.,  
Jonkers, J., Smith, J., Plumb, R.W., Taylor, R.G., Nishijima, I., Yu, Y.,  
Rogers, J., and Bradley, A.  
TITLE Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,  
JOURNAL CB10 15A, UK. http://www.sanger.ac.uk/MICR  
FEATURES  
source location/Qualifiers  
1..716  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
/clone="MHP438f23"  
/clone\_1lb="MHP"

Query Match Best Local Similarity 82.1%; Score 580.4; DB 9; Length 716;  
Matches 581; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 45 ATCTAAGTCATCAGATCATATGTCGCGGCTTTTTCGCGCTCAGTCATGCCCCAAGC 104

Db 582 ATCTAAGTCATCAGATCATATGTCGCGGCTTTTTCGCGCTCAGTCATGCCCCAAGC 104

QY 105 TGGCGCTATCTGGGAGATGGGAGAGAAAGCCCGTACCCTTTTCCCGGAGGTTGAACG 523

Db 522 TGGCGCTATCTGGGAGATGGGAGAGAAAGCCCGTACCCTTTTCCCGGAGGTTGAACG 523

QY 165 GATGAGAAAGATTGGCCAGATGACCTGCTGCTGATGAGCTTGAACGGAACGACG 463

Db 462 GATGAGAAAGATTGGCCAGATGACCTGCTGCTGATGAGCTTGAACGGAACGACG 463

QY 225 TTTCATCATGATGATGGGAGAGGTTGGCCATGACGCTTTTAAACGAGTCTGTTG 284

Db 402 TTTCATCATGATGATGGGAGAGGTTGGCCATGACGCTTTTAAACGAGTCTGTTG 284

QY 285 CAGGCCACCTGGGATACCAATTGCTGCGCGCTTTTCCGAGACAGTTCCGGATGTCACG 343

Db 342 CAGGCCACCTGGGATACCAATTGCTGCGCGCTTTTCCGAGACAGTTCCGGATGTCACG 343

QY 345 CCGAAGCGCATCAGCAACCCGAAATACCGGCGACAGCCGGAATGCGCGTGGTG 283

Db 282 CCGAAGCGCATCAGCAACCCGAAATACCGGCGACAGCCGGAATGCGCGTGGTG 283

QY 405 CAGATTAAATGACACGGTGGCGCTGGGATATTACGTACAGGAGAGGGTATCCTGGC 464

Db 222 CAGATTAAATGACACGGTGGCGCTGGGATATTACGTACAGGAGAGGGTATCCTGGC 163

QY 465 TGGATGCCGAGAAATGACATGATATACCCCGTGAATTACCCGGGGGGCGCGCTCTGTC 163

Db 162 TGGATGCCGAGAAATGACATGATATACCCCGTGAATTACCCGGGGGGCGCGCTCTGTC 524

QY 525 ATTACGTTTTTGAACCCGTGAGAGAGCGGAGACTCGCGGTGCAATGTGTTTTACAGC 103

Db 102 ATTACGTTTTTGAACCCGTGAGAGAGCGGAGACTCGCGGTGCAATGTGTTTTACAGC 584

QY 585 GTGATGAGCAGATGAAGATGCTCGACACGCTGCAAGACAGC 626

Db 42 GTGATGAGCAGATGAAGATGCTCGACACGCTGCAAGACAGC 1

RESULT 9  
CRI08810/c  
LOCUS  
DEFINITION Forward strand read from insert in 3'HPRT insertion targeting and  
ACCESSION CRI08810  
VERSION CRI08810.1 GI:49856225  
KEYWORDS GSS; genome survey sequence; MICR.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
REFERENCE Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
AUTHORS 1 (bases 1 to 786)  
Adam, D.J., Biggs, P.J., Cox, A.V., Davies, R.M., van der Weyden, L.,  
Jonkers, J., Smith, J., Plumb, R.W., Taylor, R.G., Nishijima, I., Yu, Y.,  
Rogers, J., and Bradley, A.  
TITLE Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,  
JOURNAL CB10 15A, UK. http://www.sanger.ac.uk/MICR  
FEATURES  
source location/Qualifiers  
1..786  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
/clone="MHP22016"  
/clone\_1lb="MHP"

Query Match Best Local Similarity 82.0%; Score 580; DB 9; Length 786;  
Matches 580; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 52 TGCATCAGATCATATGTCGCGGCTTTTTCGCGCTCAGTCATGCCCCAAGCTGGCGCT 111

Db 786 TGCATCAGATCATATGTCGCGGCTTTTTCGCGCTCAGTCATGCCCCAAGCTGGCGCT 111

QY 112 ATCGGGGATCGGGAGAGAAAGCCCGTACCCTTTTCCCGGAGGTTGAACGCGCATGGA 171

Db 726 ATCGGGGATCGGGAGAGAAAGCCCGTACCCTTTTCCCGGAGGTTGAACGCGCATGGA 171

QY 172 AAGAGTTTCCGAGATGATCTGCTGCTGATGACGTTGAACGGAACGACGTTTACA 667

Db 666 AAGAGTTTCCGAGATGATCTGCTGCTGATGACGTTGAACGGAACGACGTTTACA 667

QY 232 TGATGATCGGGAGAGGTTGGCCATGACGCTTTTAAACGTTGAACGTTGTTACAGGCA 607

Db 606 TGATGATCGGGAGAGGTTGGCCATGACGCTTTTAAACGTTGAACGTTGTTACAGGCA 607

QY 292 CCGGAGATACCAATTGCTGCGCGCTTTTCCGAGACAGTTCCGGATGTCACGCGCA 547

Db 546 CCGGAGATACCAATTGCTGCGCGCTTTTCCGAGACAGTTCCGGATGTCACGCGCA 547

QY 352 GCATCAGCAACCCGAAATACCGGCGACAGCCGGAATGCGCGTGGTGAGATTA 411

Db 486 GCATCAGCAACCCGAAATACCGGCGACAGCCGGAATGCGCGTGGTGAGATTA 411

QY 412 ATGACAGCGTGGCGCTGGGATATTACGTACAGGAGAGGGTATCCTGGATGTC 471

Db 426 ATGACGCGGTGGCGGCTGGGATATTAAGTCAGGAGAGACGGGATTCCTGGCTGATGC 367

Qy 472 CGCAGAAATGACATGATATACCCCTGAGTTACCCGCGGCGGCGCCTCTGTTCAATCAG 531

Db 366 CGCAGAAATGACATGATATACCCCTGAGTTACCCGCGGCGGCGCCTCTGTTCAATCAG 307

Qy 532 TTTTGAACCCGTGGAGAGACGGGACAGCTCGCGGTCGAAATGTGTTTAAAGCGGATGAG 591

Db 306 TTTTGAACCCGTGGAGAGACGGGACAGCTCGCGGTCGAAATGTGTTTAAAGCGGATGAG 247

Qy 592 AGCAGATGAGATGCTCGACACGCTGACAGAACACGACAGCT 631

Db 246 AGCAGATGAGATGCTCGACACGCTGACAGAACACGACAGCT 207

RESULT 10  
CD350776/c 691 bp mRNA linear EST 09-JUL-2003  
LOCUS UI-M-G10-csh-9-19-0-UI.r1 NIH\_BMAP\_G10 Mus musculus cDNA clone  
DEFINITION IMAGE:6853388 5', mRNA sequence.  
ACCESSION CD350776 GI:31142363  
VERSION CD350776.1  
KEYWORDS EST.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 691)  
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgebs-remail.nih.gov  
Tissue Procurement: Dr. Jim Jin, University of Iowa  
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa  
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Distribution information can be found at  
http://genome.uiowa.edu/distribution/mouse1.html  
This clone was contributed by the Brain Molecular Anatomy Project (BMAP)  
Seq primer: PYX-5.

# FEATURES

source location/Qualifiers  
1..691  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6"  
/db\_xref="taxon:10090"  
/clone="IMAGE:6853388"  
/tissue\_type="whole brain"  
/dev\_stage="embryo 13.5,14.5,16.5,17.5dpc"  
/lab\_host="DH10B (T1 phage resistant)"  
/clone\_lib="NIH BMAP\_G10"  
/note="Organ: Brain; Vector: pYX-Asc; Site 1: EcoR I; Site 2: Not I; The library was constructed according to Bonaïdo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is AGCAGACAG. This library was created for the University Iowa Brain Anatomy Project (BMAP). Gene Discovery in the Developing Mouse Nervous System", supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

ORIGIN  
Query Match 80.0%; Score 565.4; DB 6; Length 691;  
Best Local Similarity 99.8%; Pred. No. 2,3e-158;  
Matches 566; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 45 ATCTAGCTGCATCAGATCATATGCTGGGCTCTTTTCCGGCTCAGTCATCGCCCAAGC 104

Db 567 ATTCAGCTGCATCAGATCATATGCTGGGCTCTTTTCCGGCTCAGTCATCGCCCAAGC 508

Qy 105 TGGCGCTATCTGGGCGATCGGGAGAGAACCCGTGCTTTTCCGGAGGTTGAAGC 164

Db 507 TGGCGCTATCTGGGCGATCGGGAGAGAACCCGTGCTTTTCCGGAGGTTGAAGC 448

Qy 165 GCATGGAAGAAGTTTGGCCGAGAGTACTGCTGTCATTTGACGTTGACGAAACGACAG 224

Db 447 GCATGGAAGAAGTTTGGCCGAGAGTACTGCTGTCATTTGACGTTGACGAAACGACAG 388

Qy 225 TTACCATGATGATTCGGGAAGGTGTGCGCATGACGCGCTTTAAAGTGAACCTGTT 284

Db 387 TTACCATGATGATTCGGGAAGGTGTGCGCATGACGCGCTTTAAAGTGAACCTGTT 328

Qy 285 CAGGCGACCTGGGATACAGATTGCTGCGGCTTTTCCGGACACAGTTCCGGATGCTCAGC 344

Db 327 CAGGCGACCTGGGATACAGATTGCTGCGGCTTTTCCGGACACAGTTCCGGATGCTCAGC 268

Qy 345 CCGAAGCGCATCAGCAACCCGACCAATACCGGACAGACCGGAACTGCGCGGTGTG 404

Db 267 CCGAAGCGCATCAGCAACCCGACCAATACCGGACAGACCGGAACTGCGCGGTGTG 208

Qy 405 CAGATTAATGACACGCGTGGCGCTGGGATATTACGTCACGAGACGAGTATCTGCGC 464

Db 207 CAGATTAATGACACGCGTGGCGCTGGGATATTACGTCACGAGACGAGTATCTGCGC 148

Qy 465 TGGATGCGCGCAAGTATGACATGATATCCCGTAGTTACCGGCGGCGGCGCTGCTG 524

Db 147 TGGATGCGCGCAAGTATGACATGATATCCCGTAGTTACCGGCGGCGGCGCTGCTG 88

Qy 525 ATTACGCTTTTGAACCCGTGAGAGACGCGACAGCTCGGTCGAAATGTTTACAGC 584

Db 87 ATTACGCTTTTGAACCCGTGAGAGACGCGACAGCTCGGTCGAAATGTTTACAGC 28

Qy 585 GTGATGAGACATGAAAGATGCTGAC 611

Db 27 GTGATGAGACATGAAAGATGCTGAC 1

# RESULT 11

LOCUS CR035207 762 bp DNA linear GSS 05-JUN-2004  
DEFINITION Reverse strand read from insert in 5'HPRT insertion targeting and chromosome engineering clone MHPN278n13, genomic survey sequence.  
ACCESSION CR035207 GI:49768262  
VERSION GSS; genome survey sequence; MICE.  
KEYWORDS Mus musculus (house mouse)  
SOURCE Mus musculus  
ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 762)  
AUTHORS Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,L., Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y., Rogers,J. and Bradley,A.  
TITLE Direct Submission

JOURNAL Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. http://www.sanger.ac.uk/MICE

# FEATURES

source location/Qualifiers  
1..762  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
/clone="MHPN278n13"  
/clone\_lib="MHPN"

ORIGIN  
Query Match 79.8%; Score 564; DB 9; Length 762;  
Best Local Similarity 100.0%; Pred. No. 6,2e-158;  
Matches 564; Conservative 0; Mismatches 0; Indels 0; Gaps 0;





Chromosome engineering clone MHP274b23, genomic survey sequence.

ACCESSION CR139473.1 GI:49887345  
VERSION CR139473.1  
KEYWORDS GSS; genome survey sequence; MICR.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 733)  
AUTHORS Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,L.,  
Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y.,  
Rogers,J. and Bradley,A.

TITLE Direct Submission  
JOURNAL Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,  
CB10 1SA, UK. http://www.sanger.ac.uk/MICR

FEATURES  
source location/Qualifiers  
1..733  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
/clone="MHP274b23"  
/clone\_1id="MHP"

## ORIGIN

Query Match 79.5%; Score 562; DB 9; Length 733;  
Best Local Similarity 99.3%; Pred. No. 2.5e-157;  
Matches 573; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 56 TCAGGATCATATCTCGGGCTCTTTTTCGGCTCAGTATGCGCCAGACTGGCGTATCT 115  
DB 733 TCAGGATCATATCTCGGGCTCTTTTTCGGCTCAGTATGCGCCAGACTGGCGTATCT 674  
QY 116 GGGCATCGGGAGAGAAAGCCCGTCTTTTCCCGCAGGTTGAAGCGGATGGAAGA 175  
DB 673 GGGCATCGGGAGAGAAAGCCCGTCTTTTCCCGCAGGTTGAAGCGGATGGAAGA 614  
QY 176 GTTTCGCGAGGATACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 235  
DB 613 GTTTCGCGAGGATACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 554  
QY 236 GATTGCGGAAAGTGTGCGCATGCAAGCCTTTAAGGTAAGTGTGCTGCTGCTGCTGCT 295  
DB 553 GATTGCGGAAAGTGTGCGCATGCAAGCCTTTAAGGTAAGTGTGCTGCTGCTGCTGCT 494  
QY 296 GGATACCAAGTTCGTGCGCGCTTTTTCGGACAAGTTCGGATGCTGACCGCGAAGCGCAT 355  
DB 493 GGATACCAAGTTCGTGCGCGCTTTTTCGGACAAGTTCGGATGCTGACCGCGAAGCGCAT 434  
QY 356 CAGGAACCCGAAACAATACCCGCGACAGCGGAACTGCGG-TCGCGGTGTGCAATTATG 414  
DB 433 CAGGAACCCGAAACAATACCCGCGACAGCGGAACTGCGGTCGCTGTGCAATTATG 374  
QY 415 ACACGCGTGCAGCGCTGGGATATTAACGTACAGAGAGCGGGATTCCTGGCTGATGCGCG 474  
DB 373 ACACGCGTGCAGCGCTGGGATATTAACGTACAGAGAGCGGGATTCCTGGCTGATGCGCG 314  
QY 475 AGAATGAGACATGATACCCGCTGATTAACCCGCGCGCGCGCTCGTTCACTACGTTT 534  
DB 313 AGAATGAGACATGATACCCGCTGATTAACCCGCGCGCGCGCTCGTTCACTACGTTT 254  
QY 535 TTGAACCCGCTGAGAGAGCGGAGAGCTGCGCGTGCAGAAATGTGTTTACAGCTGATGAGC 594  
DB 253 TTGAACCCGCTGAGAGAGCGGAGAGCTGCGCGTGCAGAAATGTGTTTACAGCTGATGAGC 194  
QY 595 AGATGAGATGCTGCAACGCTGCAACAGCAGCT 631  
DB 193 AGATGAGATGCTGCAACGCTGCAACAGCAGCT 157

RESULT 14  
CR167678 633 bp DNA linear GSS 06-JUN-2004  
LOCUS CR167678  
DEFINITION Reverse strand read from insert in 3'HPT insertion targeting and

Chromosome engineering clone MHP161K02, genomic survey sequence.

ACCESSION CR167678  
VERSION CR167678.1 GI:49946527  
KEYWORDS GSS; genome survey sequence; MICR.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 633)  
AUTHORS Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,L.,  
Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y.,  
Rogers,J. and Bradley,A.

TITLE Direct Submission  
JOURNAL Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,  
CB10 1SA, UK. http://www.sanger.ac.uk/MICR

FEATURES  
source location/Qualifiers  
1..633  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
/clone="MHP161K02"  
/clone\_1id="MHP"

## ORIGIN

Query Match 78.5%; Score 555; DB 9; Length 633;  
Best Local Similarity 100.0%; Pred. No. 3e-155;  
Matches 555; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 77 TTTTTCGCGCTCAGTATGCGCCAGCTGCGCTATCTGGGCAATCGGGAGAGAAGC 136  
DB 633 TTTTTCGCGCTCAGTATGCGCCAGCTGCGCTATCTGGGCAATCGGGAGAGAAGC 574  
QY 137 CCGTCCCTTTTCCCGCAGGTTGAAGCGGACATGAAAGAGTTTCCGAGATGACTGCTG 196  
DB 573 CCGTCCCTTTTCCCGCAGGTTGAAGCGGACATGAAAGAGTTTCCGAGATGACTGCTG 514  
QY 197 CTGATTTGAGCTTTAGGAGAAACGACGTTTACATATGATTTGGGAAAGTGTGCGCAT 256  
DB 513 CTGATTTGAGCTTTAGGAGAAACGACGTTTACATATGATTTGGGAAAGTGTGCGCAT 454  
QY 257 GCAGCGCTTTTACCGGTGAATCTGTTGTTACAGCGCACCTGGGATACAGTTGCTGCGGCT 316  
DB 453 GCAGCGCTTTTACCGGTGAATCTGTTGTTACAGCGCACCTGGGATACAGTTGCTGCGGCT 394  
QY 317 TTTCCGAGACAGATTCGCGATGCTCAGCCCGAAGCGATGCAACCGCAATACCGG 376  
DB 393 TTTCCGAGACAGATTCGCGATGCTCAGCCCGAAGCGATGCAACCGCAATACCGG 334  
QY 377 CGACAGCCGGAACCTGCGGTGTGCAATTAATGACAGCGGTGCGGTGCGGTGCGGATA 436  
DB 333 CGACAGCCGGAACCTGCGGTGTGCAATTAATGACAGCGGTGCGGTGCGGTGCGGATA 274  
QY 437 TTACGTCAGAGAGAGCGGGATCTGCTGCTGAGTCCGCAAGAAATGACATGATACCCCG 496  
DB 273 TTACGTCAGAGAGAGCGGGATCTGCTGCTGAGTCCGCAAGAAATGACATGATACCCCG 214  
QY 497 TGAATTACCCGCGCGCGCGCTGTTCACTTCAAGTTTGTGAACCCGTGAGAGAGCGGCA 556  
DB 213 TGAATTACCCGCGCGCGCGCTGTTCACTTCAAGTTTGTGAACCCGTGAGAGAGCGGCA 154  
QY 557 GACTCGCGGTGCAATATGTTTTCAGCGTATGAGACATGAAAGTGTGCAACGCT 616  
DB 153 GACTCGCGGTGCAATATGTTTTCAGCGTATGAGACATGAAAGTGTGCAACGCT 94  
QY 617 GCAGAACCGAGCT 631  
DB 93 GCAGAACCGAGCT 79

RESULT 15  
CB520716/c 791 bp mRNA linear EST 09-JUN-2003  
LOCUS CB520716  
DEFINITION UT-M-GIO-cej-j-06-0-UT.r1 NIH\_BMAP\_GIO Mus musculus cDNA clone

IMAGE:6840391.5', mRNA sequence.  
 ACCESSION CBS20716 GI:29354071  
 VERSION C8520716.1  
 KEYWORDS EST.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 791)  
 NIH-MGC http://mgc.nhl.nih.gov/  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: c9apbs-remail.nih.gov  
 Tissue Procurement: Dr. Jim Lin, University of Iowa  
 CDNA library preparation: Dr. M. Bento Soares, University of Iowa  
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
 Clone Distribution: Distribution information can be found at  
 http://genome.uiowa.edu/distribution/mousefl.html  
 This clone was contributed by the Brain Molecular Anatomy Project  
 (BMAP)

FEATURES  
 source  
 Location/Qualifiers  
 1..791  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="C57BL/6"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:6840391"  
 /tissue\_type="whole brain"  
 /dev\_stage="embryo 13.5,14.5,16.5,17.5dpc"  
 /lab\_host="DH10B (T1 phage resistant)"  
 /clone\_id="N1H\_BMAP\_610"  
 /note="Organ: Brain; Vector: pYX-Asc; Site 1: EcoR I;  
 Site 2: Not I; The library was constructed according  
 to Bonaldi, Lennon and Soares, Genome Research, 6:791-806,  
 1996. Denatured RNA was size fractionated on a 1% agarose  
 gel. First strand cDNA synthesis was primed with oligo-dT  
 primer containing a Not I site. Double strand cDNA was  
 size selected according to mRNA size fraction, ligated  
 with EcoR I adaptor, digested with NotI and then cloned  
 directionally into pYX-Asc vector. The library tag  
 is AGCGAGACG. This library was created for the University  
 of Iowa Brain Anatomy Project (BMAP). Gene Discovery in the  
 Developing Mouse Nervous System, supported by National  
 Institute of Mental Health (NIMH), Hemlin Chin, Ph.D.,  
 program coordinator."

# ORIGIN

Query Match 78.4%; Score 554.4; DB 6; Length 791;  
 Best Local Similarity 99.6%; Pred. No. 4.8e-155;  
 Matches 566; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

```

QY 45 ATCTAGCTGATCAGATCATATCTGGGCTTTTTCGGCTCAGTCATGCCCAAGC 104
DB 568 ATCCAGCTGATCAGATCATATCTGGGCTTTTTCGGCTCAGTCATGCCCAAGC 509
QY 105 TGCGGCTATCTGGGCTATCGGAGAGAAAGCCGCTTTTTCGGAGAGTTGAACG 164
DB 508 TGCGGCTATCTGGGCTATCGGAGAGAAAGCCGCTTTTTCGGAGAGTTGAACG 449
QY 165 GCATGGAAGAGTTTCCGAGATGACTGCTGCTGATTTAGCGTGAAGCGAACG 224
DB 448 GCATGGAAGAGTTTCCGAGATGACTGCTGCTGATTTAGCGTGAAGCGAACG 389
QY 225 TTACCATATGATTTGGGAAGGTGTGGCATGACGCTTTAAAGGTGAACGTTT 284
DB 388 TTACCATATGATTTGGGAAGGTGTGGCATGACGCTTTAAAGGTGAACGTTT 329
QY 285 CAGGCCACTGGGATACGATTCGTCTGGGCTTTTCGAGACAGTTCCGATGTCAGC 344

```

```

DB 328 CAGGCCACTGGGATACGATTCGTCTGGGCTTTTCGAGACAGTTCCGATGTCAGC 269
QY 345 CCGAAGCCGATCAGCAACCCGACAAATACCGGACAGCCGAACTGCGCGGTGG 404
DB 268 CCGAAGCCGATCAGCAACCCGACAAATACCGGACAGCCGAACTGCGCGGTGG 209
QY 405 CAGATTAAATGACAGCGGCGCGCGCGGATTTAGCTCAGCGAGACGGTATCCTG 464
DB 208 CAGATTAAATGACAGCGGCGCGCGCGGATTTAGCTCAGCGAGACGGTATCCTG 149
QY 465 TGATGCCCGAGAAATGACATGATACCCGTGAGTTACCGGCGGCGCGCTCTG 524
DB 148 TGATGCCCGAGAAATGACATGATACCCGTGAGTTACCGGCGGCGCGCTCTG 89
QY 525 ATTACGTTTGTGAACCCGTTGAGAGACGGGCGAGACTCGCGGTGCAATGTGTTTAC-AG 583
DB 88 ATTACGTTTGTGAACCCGTTGAGAGACGGGCGAGACTCGCGGTGCAATGTGTTTAC-AG 29
QY 584 CGTGAATGAGACAGATGAAATGATGCTCGAC 611
DB 28 CGTGAATGAGACAGATGAAATGATGCTCGAC 1

```

Search completed: January 26, 2005, 12:13:54  
 Job time : 2228.98 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 25, 2005, 23:20:55 ; Search time 334.971 Seconds  
(without alignments)  
11079.575 Million cell updates/sec

Title: US-10-826-523-40

Perfect score: 707  
Sequence: 1 ggaaccatgcgcgaattt.....aaatgacgacatggatcc 707

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_23Sep04:\*  
1: geneseq19908:\*  
2: geneseq19908:\*  
3: geneseq20008:\*  
4: geneseq2001as:\*  
5: geneseq2001bs:\*  
6: geneseq2002as:\*  
7: geneseq2002bs:\*  
8: geneseq2003as:\*  
9: geneseq2003bs:\*  
10: geneseq2003cs:\*  
11: geneseq2003ds:\*  
12: geneseq2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	707	100.0	707	10 ACAS5353	ACAS5353 Eukaryotl
2	707	100.0	4613	10 ACAS5360	ACAS5360 Transform
3	705.4	99.8	8999	10 ACAS5361	ACAS5361 Transform
4	705.4	99.8	9009	10 ACAS5362	ACAS5362 Transform
5	705.4	99.8	9012	10 ACAS5363	ACAS5363 Transform
6	685	96.9	4943	10 ACAS5369	ACAS5369 Transform
7	685	96.9	4943	10 ACAS5367	ACAS5367 Transform
8	681.8	96.4	4944	10 ACAS5368	ACAS5368 Transform
9	585.4	82.8	48502	12 ADP70043	ADP70043 Nucleic a
10	580.6	82.1	4164	4 AAS44733	Aas44733 Bacteriop
11	566.2	80.1	26565	6 AAB78924	Aab78924 E. coli C
12	566.2	80.1	26565	6 ADH80491	Adh80491 Escherich
13	566.2	80.1	46819	9 ACD19059	Acid19059 E. coli 0
14	564.6	79.9	46897	10 ADC00585	Adc00585 Enterohae
15	561.4	79.4	22306	9 ACD19238	Acid19238 E. coli 0
16	561.4	79.4	91740	10 ADC00956	Adc00956 Enterohae
17	519.2	73.4	3661	10 ACAS5354	ACAS5354 Transform
18	511	72.3	4941	10 ACAS5366	ACAS5366 Transform
19	511	72.3	4952	10 ACAS5365	ACAS5365 Transform
20	499	70.6	4951	10 ACAS5364	ACAS5364 Transform
21	425.8	60.2	1427	6 ABQ50926	ABQ50926 Oligonuc

C 22	425.8	60.2	1427	6	ABQ50927	ABQ50927 Oligonuc
C 23	382.6	55.5	3822	5	AAS94485	Aas94485 DNA encod
C 24	382.6	54.1	1427	6	ABQ50928	ABQ50928 Oligonuc
C 25	382.6	54.1	1427	6	ABQ50929	ABQ50929 Oligonuc
C 26	223	31.5	34063	9	ACD19083	Acid19083 E. coli 0
C 27	223	31.5	44029	10	ADC00710	Adc00710 Enterohae
C 28	223	31.5	49650	10	ADC00365	Adc00365 Enterohae
C 29	222.2	31.4	134141	6	ABN83487	Abn83487 Escherich
C 30	171.2	24.2	822	5	AAS80174	Aas80174 DNA encod
C 31	96	13.6	847	2	AAZ27249	Aaz27249 Human sec
C 32	78	11.0	9980	10	ACAS5352	ACAS5352 PTAO-P/L
C 33	73.8	10.4	5194	4	AAC85595	Aac85595 Plasmid P
C 34	73.8	10.4	5194	4	ADA09845	Ada09845 PiggyBac
C 35	73.4	10.4	5194	4	AAC85596	Aac85596 Plasmid P
C 36	73.4	10.4	5194	9	ADA09846	Ada09846 PiggyBac
C 37	72.8	10.3	7670	10	ACAS5351	ACAS5351 PTAO-P/L
C 38	67.8	9.6	2480	8	ABV76195	Abv76195 Trichoplu
C 39	67.8	9.6	5679	9	AAC85602	Aac85602 Heliper pl
C 40	67.8	9.6	5679	9	ADA09857	Ada09857 Heliper pl
C 41	67.8	9.6	6723	4	AAC85593	Aac85593 phsp70/op
C 42	67.8	9.6	6723	4	AAC85592	Aac85592 Plasmid P
C 43	67.8	9.6	6723	9	ADA09843	Ada09843 PiggyBac
C 44	67.8	9.6	6723	9	ADA09842	Ada09842 Plasmid P
C 45	67.8	9.6	7560	4	AAC85599	Aac85599 Plasmid P

## ALIGNMENTS

RESULT 1	ACAS5353
ID	ACAS5353 standard; DNA, 707 BP.
AC	ACAS5353:
AC	ACAS5353:
DT	09-JUN-2003 (first entry)
XX	
DE	Eukaryotic transformation vector piggyBac ITR cartridge.
XX	
KW	PiggyBac; transposon; eukaryotic transformation vector; ds;
KW	transformed cell; transformed embryo; transgenic; ITR cartridge.
XX	
OS	Synthetic.
XX	
PN	US2002173634-A1.
XX	
PD	21-NOV-2002.
XX	
PF	30-OCT-2001; 2001US-00001189.
XX	
PR	31-OCT-2000; 2000US-0244677P.
PR	01-NOV-2000; 2000US-0244984P.
XX	
PA	(FRAS/) FRASER M J.
PA	(LTX/) LI X.
PA	(BEAM/) BEAM T.
PA	(HDAV/) HDA-VAN A.
PI	Fraser MJ, Li X, Beam T, Hua-Van A;
XX	
DR	WPI; 2003-352597/33.
XX	
PT	New DNA molecule in the transposon piggyBac, useful for transferring
PT	genes into host cells or embryos for transforming the cells of embryos
PT	that can be used in making transgenic organisms.
XX	
PS	Example 3; Fig 3(C1); 151bp; English.
XX	
CC	The invention describes a DNA molecule comprising at least 163
CC	consecutive nucleotide base pairs of the 3' terminal region beginning at
CC	the 3' terminal base pair, and at least 125 consecutive nucleotide base
CC	pairs of the 5' terminal region beginning at the 5' terminal base pair, of
CC	the piggyBac molecule. The region extends from the restriction site SacI

to the end of the piggyBac molecule. The DNA molecule in the transposon piggyBac is useful for transferring genes into host cells or embryos for transforming the cells of embryos. The transformed cells or embryos are useful for developing or making transgenic organisms. This sequence represents a minimal sequence cartridge of the eukaryotic transformation vector piggyBac

Sequence 707 BP; 164 A; 174 C; 208 G; 161 T; 0 U; 0 Other;

Query Match 100.0%; Score 707; DB 10; Length 707;  
Best Local Similarity 100.0%; Pred. No. 2,3e-213;  
Matches 707; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGATCCCATGCGTCAATTTTACGAGATATCTTTTACGGTTATCTAGCTGCATCAGG 60  
DB 1 GGATCCCATGCGTCAATTTTACGAGATATCTTTTACGGTTATCTAGCTGCATCAGG 60  
QY 61 ATCATATGCTGGGCTCTTTTCCGGCTCACTCACTCCCAAGCTGGCGTATCTGGCA 120  
DB 61 ATCATATGCTGGGCTCTTTTCCGGCTCACTCACTCCCAAGCTGGCGTATCTGGCA 120  
QY 121 TCGGGAG 180  
DB 121 TCGGGAG 180  
QY 181 CCGAGATGACTGCTGCTGATTAAGCTTGAAGGAAAGCAAGTTTACATGATGATTC 240  
DB 181 CCGAGATGACTGCTGCTGATTAAGCTTGAAGGAAAGCAAGTTTACATGATGATTC 240  
QY 241 GGGAGAGTGTGGCCATGACAGCTTTTAAAGCTGTTAGCTGTTAGCTGTTAGCT 300  
DB 241 GGGAGAGTGTGGCCATGACAGCTTTTAAAGCTGTTAGCTGTTAGCTGTTAGCT 300  
QY 301 CCGATTCGTGCGCGCTTTTCCGGCAACAGTTCCGATGATGATGATGATGATGAT 360  
DB 301 CCGATTCGTGCGCGCTTTTCCGGCAACAGTTCCGATGATGATGATGATGATGAT 360  
QY 361 ACCCGAACAATACCGCGACAGCCGGAATCTGCGCTGCTGCTGCTGCTGCTGCT 420  
DB 361 ACCCGAACAATACCGCGACAGCCGGAATCTGCGCTGCTGCTGCTGCTGCTGCT 420  
QY 421 GTGCGCGCTGCGATTTTACGTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480  
DB 421 GTGCGCGCTGCGATTTTACGTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480  
QY 481 GGACATGATACCCGCTGATGTTACCGCGCGCGCGCGCTGCTGCTGCTGCTGCT 540  
DB 481 GGACATGATACCCGCTGATGTTACCGCGCGCGCGCGCTGCTGCTGCTGCTGCT 540  
QY 541 CCGTGGAG 600  
DB 541 CCGTGGAG 600  
QY 601 AGATGCTGACACGCTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660  
DB 601 AGATGCTGACACGCTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660  
QY 661 TGAAGTACGTTAAAGATTAATCATGCGTAAATTAAGAGAGAGAGAGAGAGAG 707  
DB 661 TGAAGTACGTTAAAGATTAATCATGCGTAAATTAAGAGAGAGAGAGAGAGAG 707

RESULT 2  
ACAS5360  
ID ACAS5360 standard; DNA; 4613 BP.  
XX  
AC ACAS5360;  
XX  
XX  
DT 06-JUN-2003 (first entry)  
XX  
XX  
DE Transformation vector piggyBac related plasmid pCR11-1TR.  
XX  
XX PiggyBac; transposon; eukaryotic transformation vector; ds;

KW transformed cell; transformed embryo; transgenic; plasmid; cyclic;  
XX circular.  
XX  
XX  
OS Synthetic.

US2002173634-A1.

21-NOV-2002.

30-OCT-2001; 2001US-00001189.

31-OCT-2000; 2000US-0244677P.

01-NOV-2000; 2000US-0244984P.

(FRAS/) FRASER M J.

(LIXX/) LI X.

(BEAM/) BEAM T.

(HUA/V/) HUA-VAN A.

Fraser MJ, Li X, Beam T, Hua-Van A;

WPI; 2003-352597/33.

P-PSDB; ABU70357, ABU70358.

New DNA molecule in the transposon piggyBac, useful for transferring genes into host cells or embryos for transforming the cells of embryos that can be used in making transgenic organisms.

Disclosure; Fig 10(B); 151pp; English.

The invention describes a DNA molecule comprising at least 163 consecutive nucleotide base pairs of the 3' terminal region beginning at the 3' terminal base pair, and at least 125 consecutive nucleotide base pairs of the 5' terminal region beginning at the 5' terminal base pair, of the piggyBac molecule. The region extends from the restriction site SacI to the end of the piggyBac molecule. The DNA molecule in the transposon piggyBac is useful for transferring genes into host cells or embryos for transforming the cells of embryos. The transformed cells or embryos are useful for developing or making transgenic organisms. This sequence represents plasmid used in the creation of minimal sequence eukaryotic transformation vector piggyBac

Sequence 4613 BP; 1064 A; 1183 C; 1273 G; 1093 T; 0 U; 0 Other;

Query Match 100.0%; Score 707; DB 10; Length 4613;  
Best Local Similarity 100.0%; Pred. No. 6e-213;  
Matches 707; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGATCCCATGCGTCAATTTTACGAGATATCTTTTACGGTTATCTAGCTGCATCAGG 60  
DB 294 GGATCCCATGCGTCAATTTTACGAGATATCTTTTACGGTTATCTAGCTGCATCAGG 353  
QY 61 ATCATATGCTGGGCTCTTTTCCGGCTCACTCACTCCCAAGCTGGCGTATCTGGCA 120  
DB 354 ATCATATGCTGGGCTCTTTTCCGGCTCACTCACTCCCAAGCTGGCGTATCTGGCA 413  
QY 121 TCGGGAG 180  
DB 414 TCGGGAG 473  
QY 181 CCGAGATGACTGCTGCTGATTAAGCTTGAAGGAAAGCAAGTTTACATGATGATTC 240  
DB 474 CCGAGATGACTGCTGCTGATTAAGCTTGAAGGAAAGCAAGTTTACATGATGATTC 533  
QY 241 GGGAGAGTGTGGCCATGACAGCTTTTAAAGCTGTTAGCTGTTAGCTGTTAGCT 300  
DB 534 GGGAGAGTGTGGCCATGACAGCTTTTAAAGCTGTTAGCTGTTAGCTGTTAGCT 593  
QY 301 CCGATTCGTGCGCGCTTTTCCGGCAACAGTTCCGATGATGATGATGATGATGAT 360  
DB 594 CCGATTCGTGCGCGCTTTTCCGGCAACAGTTCCGATGATGATGATGATGATGAT 653  
QY 361 ACCCGAACAATACCGCGACAGCCGGAATCTGCGCTGCTGCTGCTGCTGCTGCT 420

DB 654 ACCGGAACATACCGGCGACACCGGAACTCCCGCGGTGTGAGATTATGACGCG 713  
 QY 421 GTGCGCGCTGGATATTACCTGACGAGACGCGGTATCTGGTGTGATGCCGAGAAAT 480  
 DB 714 GTGCGCGCTGGATATTACCTGACGAGACGCGGTATCTGGTGTGATGCCGAGAAAT 773  
 QY 481 GGACATGATATCCCGTGAATTAACCGCGCGCGCTCTGTTCACTTACGTTTTTGAAC 540  
 DB 774 GGACATGATATCCCGTGAATTAACCGCGCGCGCTCTGTTCACTTACGTTTTTGAAC 833  
 QY 541 CCGTGAAGACGCGGACAGCTCGCGGTCAAAATGTGTTTAAAGCGTGAAGACAGATGA 600  
 DB 834 CCGTGAAGACGCGGACAGCTCGCGGTCAAAATGTGTTTAAAGCGTGAAGACAGATGA 893  
 QY 601 AGATGCTGACACGCTGCAAGAACACGAGCTAGATTAACTTAAAGATTAATCATATTG 660  
 DB 894 AGATGCTGACACGCTGCAAGAACACGAGCTAGATTAACTTAAAGATTAATCATATTG 953  
 QY 661 TGACGTACGTTAAAGATTAATCATGCTTAAATTGACGATGGATCC 707  
 DB 954 TGACGTACGTTAAAGATTAATCATGCTTAAATTGACGATGGATCC 1000

RESULT 3  
 ACAS5361/C  
 ID ACAS5361 standard; DNA; 8999 BP.  
 XX ACAS5361;  
 AC ACAS5361;  
 DT 06-JUN-2003 (first entry)  
 XX

DE Transformation vector piggyBAC related plasmid p(PZ)-Bac-EYFP.

XX PiggyBac; transposon; eukaryotic transformation vector; ds;  
 KM transformed cell; transformed embryo; transgenic; plasmid; cyclic;  
 KM circular.

XX Synthetic.

XX US202173634-A1.

XX 21-NOV-2002.

XX 30-OCT-2001; 2001US-00001189.

XX 31-OCT-2000; 2000US-0244677P.

XX 01-NOV-2000; 2000US-0244984P.

XX (FRASER M J.

XX (LIXX/) LI X.

XX (BEAM/) BEAM T.

XX (HUA/V/) HUA-VAN A.

XX Fraser MJ, Li X, Beam T, Hua-Van A;

XX WPI; 2003-352597/33.

XX

XX Example 11; Fig 12(B); 151pp; English.

CC The invention describes a DNA molecule comprising at least 163  
 CC consecutive nucleotide base pairs of the 3' terminal region beginning at  
 CC the 3' terminal base pair, and at least 125 consecutive nucleotide base  
 CC pairs of the 5' terminal region beginning at the 5' terminal base pair, of  
 CC the piggyBac molecule. The region extends from the restriction site SacI  
 CC to the end of the piggyBac molecule. The DNA molecule in the transposon  
 CC piggyBac is useful for transferring genes into host cells or embryos for  
 CC transforming the cells or embryos. The transformed cells or embryos are  
 CC useful for developing or making transgenic organisms. This sequence

CC represents plasmid used in the creation of minimal sequence eukaryotic  
 CC transformation vector piggyBac

SO Sequence 8999 BP; 2180 A; 2416 C; 2242 G; 2061 T; 0 U; 0 Other;

Query Match 99.8%; Score 705.4; DB 10; Length 8999;

Best Local Similarity 99.9%; Pred. No. 2,76-212;

Matches 706; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGATCCCATGCGTCAATTTTAAAGCAGACTATCTTTCTAGAGTTAATCTAGCTGATCAGG 60  
 DB 8950 GGATCCCATGCGTCAATTTTAAAGCAGACTATCTTTCTAGAGTTAATCTAGCTGATCAGG 8891  
 QY 61 ATCATATTCGTCGGCTTTTTCGCGCTCACTATCCGCCAAGCTGCGCTATCTGGGCA 120  
 DB 8890 ATCATATTCGTCGGCTTTTTCGCGCTCACTATCCGCCAAGCTGCGCTATCTGGGCA 8831  
 QY 121 TCGGGAGGAAGAAAGCCGTCCTTTTCCGCGAGTTGAAGCGCATGGAAGAGTTTG 180  
 DB 8830 TCGGGAGGAAGAAAGCCGTCCTTTTCCGCGAGTTGAAGCGCATGGAAGAGTTTG 8771  
 QY 181 CCGAGATGACTGCTGCTGATTTGACGTTGACGAAACCGCATTTACATGATTC 240  
 DB 8770 CCGAGATGACTGCTGCTGATTTGACGTTGACGAAACCGCATTTACATGATTC 8711  
 QY 241 GGGAAAGTGTGCGCATGACGCTTTTAAAGCTGAATCTTCTGCTGATGCGGCACTGGGATA 300  
 DB 8710 GGGAAAGTGTGCGCATGACGCTTTTAAAGCTGAATCTTCTGCTGATGCGGCACTGGGATA 8651  
 QY 301 CCAAGTTCGCGCGCTTTTCCGAGACAGATTCGCGAGTGTGACGCGGAGCGCATCGCA 360  
 DB 8650 CCAAGTTCGCGCGCTTTTCCGAGACAGATTCGCGAGTGTGACGCGGAGCGCATCGCA 8591  
 QY 361 ACCGGAACAAATACCGGAGACAGCCGAACTGCGTCCGCTGTGTCAGATTAAATGACAGCG 420  
 DB 8590 ACCGGAACAAATACCGGAGACAGCCGAACTGCGTCCGCTGTGTCAGATTAAATGACAGCG 8531  
 QY 421 GTGCGCGCTGGATATTACCTGACGAGACGCGGTATCTGCTGATGCGGCAAAAT 480  
 DB 8530 GTGCGCGCTGGATATTACCTGACGAGACGCGGTATCTGCTGATGCGGCAAAAT 8471  
 QY 481 GGACATGATATCCCGTGAATTAACCGCGCGCGCTCTGTTCACTTACGTTTTTGAAC 540  
 DB 8470 GGACATGATATCCCGTGAATTAACCGCGCGCGCTCTGTTCACTTACGTTTTTGAAC 8411  
 QY 541 CCGTGAAGACGCGGACAGCTCGCGGTCAAAATGTGTTTAAAGCGTGAAGACAGATGA 600  
 DB 8410 CCGTGAAGACGCGGACAGCTCGCGGTCAAAATGTGTTTAAAGCGTGAAGACAGATGA 8351  
 QY 601 AGATGCTGACACGCTGCAAGAACACGAGCTAGATTAACTTAAAGATTAATCATATTG 660  
 DB 8350 AGATGCTGACACGCTGCAAGAACACGAGCTAGATTAACTTAAAGATTAATCATATTG 8291  
 QY 661 TGACGTACGTTAAAGATTAATCATGCTTAAATTGACGATGGATCC 707  
 DB 8290 TGACGTACGTTAAAGATTAATCATGCTTAAATTGACGATGGATCC 8244

RESULT 4  
 ACAS5362/C  
 ID ACAS5362 standard; DNA; 9009 BP.  
 XX ACAS5362;  
 AC ACAS5362;  
 DT 06-JUN-2003 (first entry)  
 XX  
 DE Transformation vector piggyBAC related plasmid p(PZ)-Bac-EYFP.  
 XX PiggyBac; transposon; eukaryotic transformation vector; ds;  
 KM transformed cell; transformed embryo; transgenic; plasmid; cyclic;  
 KM circular.  
 XX Synthetic.

XX US2002173634-A1.  
 XX 21-NOV-2002.  
 XX 30-OCT-2001; 2001US-00001189.  
 XX 31-OCT-2000; 2000US-0244677P.  
 XX 01-NOV-2000; 2000US-0244984P.  
 XX (FRAS/) FRASER M J.  
 XX (LIXX/) LI X.  
 XX (BEAM/) BEAM T.  
 XX (HUA/V/) HUA-VAN A.  
 XX PI  
 XX Fraser MJ, Li X, Beam T, Hua-Van A;  
 XX WPI; 2003-352597/33.  
 XX  
 XX New DNA molecule in the transposon piggyBac, useful for transferring  
 XX genes into host cells or embryos for transforming the cells of embryos  
 XX that can be used in making transgenic organisms.  
 XX  
 XX Example 11; Fig 13(B); 151pp; English.  
 XX  
 XX The invention describes a DNA molecule comprising at least 163  
 XX consecutive nucleotide base pairs of the 3' terminal region beginning at  
 XX the 3' terminal base pair, and at least 125 consecutive nucleotide base  
 XX pairs of the 5' terminal region beginning at the 5' terminal base pair, of  
 XX the piggyBac molecule. The region extends from the restriction site SacI  
 XX to the end of the piggyBac molecule. The DNA molecule in the transposon  
 XX piggyBac is useful for transferring genes into host cells or embryos for  
 XX transforming the cells of embryos. The transformed cells or embryos are  
 XX useful for developing or making transgenic organisms. This sequence  
 XX represents plasmid used in the creation of minimal sequence eukaryotic  
 XX transformation vector piggyBac  
 XX  
 XX Sequence 9009 BP; 2179 A; 2429 C; 2341 G; 2058 T; 0 U; 2 Other;  
 Query Match 99.8%; Score 705.4; DB 10; Length 9009;  
 Best Local Similarity 99.9%; Pred. No. 2.7e-212;  
 Matches 706; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGATCCATGCGTCAATTTTACGACACTATCTTTCTAGAGTTATCTAGCTGACAG 60  
 DB 8960 GGATCCATGCGTCAATTTTACGACACTATCTTTCTAGAGTTATCTAGCTGACAG 60  
 QY 61 ATCATATCGTGGGCTCTTTTCCGGCTCAGTCAATCCCAAGCTGCGCTATCTGGGCA 120  
 DB 8900 ATCATATCGTGGGCTCTTTTCCGGCTCAGTCAATCCCAAGCTGCGCTATCTGGGCA 120  
 QY 121 TCGGGGAGGAAAGAGCCCGTCCCTTTTCCGCGAGTTGAAGCGGATGGAAAGATTG 180  
 DB 8840 TCGGGGAGGAAAGAGCCCGTCCCTTTTCCGCGAGTTGAAGCGGATGGAAAGATTG 180  
 QY 181 CCGAGATGATGCTGCTGCTGATGAGTGAAGGAAAGCAAGCTTTACATGATGATTC 240  
 DB 8780 CCGAGATGATGCTGCTGCTGATGAGTGAAGGAAAGCAAGCTTTACATGATGATTC 240  
 QY 241 GGGAAAGTGGGCGCATGACAGCCTTTTAAACGCTGAATGTTTCGTTACAGGCGACATGGGATA 300  
 DB 8720 GGGAAAGTGGGCGCATGACAGCCTTTTAAACGCTGAATGTTTCGTTACAGGCGACATGGGATA 300  
 QY 301 CCAAGTTGTCGCGGCTTTTCCGGAACAATTCGAGTGTGACCCGGAAGGCAATCAGCA 360  
 DB 8660 CCAAGTTGTCGCGGCTTTTCCGGAACAATTCGAGTGTGACCCGGAAGGCAATCAGCA 360  
 QY 361 ACCCGAACAATACCGCGCAAGCCGGAACCTGCGCGGATGAGATTAATGACGCG 420  
 DB 8600 ACCCGAACAATACCGCGCAAGCCGGAACCTGCGCGGATGAGATTAATGACGCG 420  
 QY 421 GTGCGGCGTGGGATATTAATGATGATGAGGATGCTGCTGATGATGCGGCAAGAT 480  
 DB 421 GTGCGGCGTGGGATATTAATGATGATGAGGATGCTGCTGATGATGCGGCAAGAT 480

DB 8540 GTGCGGCGCTGGGATATTAATGATGATGAGGATGCTGCTGATGATGCGGCAAGAT 8481  
 QY 481 GGACATGATATCCCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540  
 DB 8480 GGACATGATATCCCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540  
 QY 541 CCGTGAAGGAGCGGCGAGACTCGCGGTGCAAAATGTTTAAACGCGGATGAGGAGATGA 600  
 DB 8420 CCGTGAAGGAGCGGCGAGACTCGCGGTGCAAAATGTTTAAACGCGGATGAGGAGATGA 600  
 QY 601 AGATGCTGCAACGCTGCAAGAAACAGCAGCTGATTAACCTTAAGAAATATATATTTG 660  
 DB 8360 AGATGCTGCAACGCTGCAAGAAACAGCAGCTGATTAACCTTAAGAAATATATATTTG 660  
 QY 661 TGACGTACGTTAAAGATTAATGATGCTGAATTAATGATGATGATGATGATGATGATGAT 707  
 DB 8300 TGACGTACGTTAAAGATTAATGATGCTGAATTAATGATGATGATGATGATGATGATGAT 8254

RESULT 5  
 ACAS5363/c  
 ID ACAS5363 standard; DNA; 9012 BP.  
 XX ACAS5363;  
 AC  
 XX 06-JUN-2003 (first entry)  
 XX  
 XX Transformation vector piggyBac related plasmid p(pZ)-Bac-BGFP.  
 XX  
 XX piggyBac; transposon; eukaryotic transformation vector; ds;  
 XX circular.  
 XX  
 XX Synthetic.  
 XX  
 XX US2002173634-A1.  
 XX  
 XX 21-NOV-2002.  
 XX  
 XX 30-OCT-2001; 2001US-00001189.  
 XX  
 XX 31-OCT-2000; 2000US-0244677P.  
 XX  
 XX 01-NOV-2000; 2000US-0244984P.  
 XX  
 XX (FRAS/) FRASER M J.  
 XX (LIXX/) LI X.  
 XX (BEAM/) BEAM T.  
 XX (HUA/V/) HUA-VAN A.  
 XX  
 XX Fraser MJ, Li X, Beam T, Hua-Van A;  
 XX WPI; 2003-352597/33.  
 XX  
 XX New DNA molecule in the transposon piggyBac, useful for transferring  
 XX genes into host cells or embryos for transforming the cells of embryos  
 XX that can be used in making transgenic organisms.  
 XX  
 XX Example 11; Fig 14(B); 151pp; English.  
 XX  
 XX The invention describes a DNA molecule comprising at least 163  
 XX consecutive nucleotide base pairs of the 3' terminal region beginning at  
 XX the 3' terminal base pair, and at least 125 consecutive nucleotide base  
 XX pairs of the 5' terminal region beginning at the 5' terminal base pair, of  
 XX the piggyBac molecule. The region extends from the restriction site SacI  
 XX to the end of the piggyBac molecule. The DNA molecule in the transposon  
 XX piggyBac is useful for transferring genes into host cells or embryos for  
 XX transforming the cells of embryos. The transformed cells or embryos are  
 XX useful for developing or making transgenic organisms. This sequence  
 XX represents plasmid used in the creation of minimal sequence eukaryotic  
 XX transformation vector piggyBac  
 XX  
 XX Sequence 9012 BP; 2180 A; 2419 C; 2353 G; 2060 T; 0 U; 0 Other;



Query Match 99.8%; Score 705.4; DB 10; Length 9012;  
 Best Local Similarity 99.9%; Pred. No. 2.7e-212;  
 Matches 706; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 GGATCCATGCGTCAATTTTACGACAGATCTTTCTAGGGTTAACTAGTCATCAGG 60  
 DB 8963 GGATCCATGCGTCAATTTTACGACAGATCTTTCTAGGGTTAACTAGTCATCAGG 8904

61 ATCATATGCTGGGCTTTTTCGGCTCAGTCAATCCGCAAGCTGGCGCTATCTGGGCA 120  
 DB 8903 ATCATATGCTGGGCTTTTTCGGCTCAGTCAATCCGCAAGCTGGCGCTATCTGGGCA 8844

121 TCGGGAGAGAAAGACCCGTCCTTTCCCGGAGGTTGAAGCGGCATGAAAGATTG 180  
 DB 8843 TCGGGAGAGAAAGACCCGTCCTTTCCCGGAGGTTGAAGCGGCATGAAAGATTG 8784

181 CCGAGATGACTGCTGCTGATTTGACGTGACGAAACGACGTTTACATGATGATTC 240  
 DB 8783 CCGAGATGACTGCTGCTGATTTGACGTGACGAAACGACGTTTACATGATGATTC 8724

241 GGAAGGTGTGGCCATGACACGCTTTAAAGGTGAACCTGTTCAAGCCACCTGGATA 300  
 DB 8723 GGAAGGTGTGGCCATGACACGCTTTAAAGGTGAACCTGTTCAAGCCACCTGGATA 8664

301 CCAAGTTCGTCGGGCTTTTCCGGACACAGTTCCGATGGTGAAGCCGAAAGGCATCAGCA 360  
 DB 8663 CCAAGTTCGTCGGGCTTTTCCGGACACAGTTCCGATGGTGAAGCCGAAAGGCATCAGCA 8604

361 ACCGAAACAATACCGGCGACAGCCGGAACCTGCGCGGTGACGATTAATGACAGCG 420  
 DB 8603 ACCGAAACAATACCGGCGACAGCCGGAACCTGCGCGGTGACGATTAATGACAGCG 8544

421 GTGCGGCGCTGGGATTTAAGTCAAGCAGAGAGCGGTATCTGCGTGAATCCGCAAAAT 480  
 DB 8543 GTGCGGCGCTGGGATTTAAGTCAAGCAGAGAGCGGTATCTGCGTGAATCCGCAAAAT 8484

481 GGACATGATACCCCGTGAAGTACCGGCGGCGGCGCTCGTTCACTTACAGTTTGAAC 540  
 DB 8483 GGACATGATACCCCGTGAAGTACCGGCGGCGGCGCTCGTTCACTTACAGTTTGAAC 8424

541 CCGTGAAGACGCGGACAGCTCGCGGTGCAAAATGTTTTACAGCTGATGAGCAGATGA 600  
 DB 8423 CCGTGAAGACGCGGACAGCTCGCGGTGCAAAATGTTTTACAGCTGATGAGCAGATGA 8364

601 AGATCTTGAACCGCTGCAAAACAGCAGCTAGATTAACTGAAAGATTAATCATATTG 660  
 DB 8363 AGATCTTGAACCGCTGCAAAACAGCAGCTAGATTAACTGAAAGATTAATCATATTG 8304

661 TGACGTACGTTAAAGATTAATCATGCGTAAATTTGACGATGGGATCC 707  
 DB 8303 TGACGTACGTTAAAGATTAATCATGCGTAAATTTGACGATGGGATCC 8257

RESULT 6  
 ID ACAS5369  
 ACAS5369 standard; DNA: 4943 BP.  
 ACAS5369;  
 06-JUN-2003 (first entry)  
 Transformation vector piggyBAC related plasmid pBS-ITR-EYFP.  
 PiggyBac; transposon; eukaryotic transformation vector; ds;  
 transformed cell; transformed embryo; transgenic; plasmid; cyclic;  
 circular.  
 OS Synthetic.  
 XX  
 XX US2002173634-A1.  
 XX  
 XX PD 21-NOV-2002.  
 XX

PF 30-OCT-2001; 2001US-00001189.  
 XX  
 PR 31-OCT-2000; 2000US-0244677P.  
 PR 01-NOV-2000; 2000US-0244984P.  
 XX  
 XX (FRAS/) FRASER M J.  
 PA (LIXX/) LI X.  
 PA (BEAM/) BEAM T.  
 PA (HDAV/) HDA-VAN A.  
 XX  
 PI Fraser MJ, Li X, Beam T, Hua-Van A;  
 DR WPI, 2003-352597/33.  
 XX  
 XX New DNA molecule in the transposon piggyBac, useful for transferring  
 PT genes into host cells or embryos for transforming the cells of embryos  
 PT that can be used in making transgenic organisms.  
 XX  
 PS Example 6; Fig 20(B); 151pp; English.  
 XX  
 XX The invention describes a DNA molecule comprising at least 163  
 CC consecutive nucleotide base pairs of the 3' terminal region beginning at  
 CC the 3' terminal base pair, and at least 125 consecutive nucleotide base  
 CC pairs of the 5' terminal region beginning at the 5' terminal base pair, of  
 CC the piggyBac molecule. The region extends from the restriction site SacI  
 CC to the end of the piggyBac molecule. The DNA molecule in the transposon  
 CC piggyBac is useful for transferring genes into host cells or embryos for  
 CC transforming the cells of embryos. The transformed cells or embryos are  
 CC useful for developing or making transgenic organisms. This sequence  
 CC represents plasmid used in the creation of minimal sequence eukaryotic  
 CC transformation vector piggyBac  
 XX

Sequence 4943 BP; 1220 A; 1302 C; 1245 G; 1176 T; 0 U; 0 Other;  
 SQ

Query Match 96.9%; Score 685; DB 10; Length 4943;  
 Best Local Similarity 99.2%; Pred. No. 5.9e-206;  
 Matches 701; Conservative 0; Mismatches 0; Indels 6; Gaps 1;

1 GGATCCATGCGTCAATTTTACGACAGATCTTTCTAGGGTTAACTAGTCATCAGG 60  
 DB 719 GGATCCATGCGTCAATTTTACGACAGATCTTTCTAGGGTTAACTAGTCATCAGG 778

61 ATCATATGCTGGGCTTTTTCGGCTCAGTCAATCCGCAAGCTGGCGCTATCTGGGCA 120  
 DB 779 ATCATATGCTGGGCTTTTTCGGCTCAGTCAATCCGCAAGCTGGCGCTATCTGGGCA 838

121 TCGGGAGAGAAAGACCCGTCCTTTCCCGGAGGTTGAAGCGGCATGAAAGATTG 180  
 DB 839 TCGGGAGAGAAAGACCCGTCCTTTCCCGGAGGTTGAAGCGGCATGAAAGATTG 898

181 CCGAGATGACTGCTGCTGATTTGACGTTGAGCAAAACGACGTTTACATGATGATTC 240  
 DB 899 CCGAGATGACTGCTGCTGATTTGACGTTGAGCAAAACGACGTTTACATGATGATTC 958

241 GGAAGGTGTGGCCATGACACGCTTTAAAGGTGAACCTGTTCAAGCCACCTGGATA 300  
 DB 959 GGAAGGTGTGGCCATGACACGCTTTAAAGGTGAACCTGTTCAAGCCACCTGGATA 1018

301 CCAAGTTCGTCGGGCTTTTCCGGACACAGTTCCGATGGTGAAGCCGAAAGGCATCAGCA 360  
 DB 1019 CCAAGTTCGTCGGGCTTTTCCGGACACAGTTCCGATGGTGAAGCCGAAAGGCATCAGCA 1078

361 ACCGAAACAATACCGGCGACAGCCGGAACCTGCGCGGTGACGATTAATGACAGCG 420  
 DB 1079 ACCGAAACAATACCGGCGACAGCCGGAACCTGCGCGGTGACGATTAATGACAGCG 1138

421 GTGCGGCGCTGGGATTTAAGTCAAGCAGAGCGGTTTCCGCTGATGCGGAGAAAT 480  
 DB 1139 GTGCGGCGCTGGGATTTAAGTCAAGCAGAGCGGTTTCCGCTGATGCGGAGAAAT 1198

481 GGACATGATACCCCGTGAAGTAACTGCGGCGGCGCTCGTTCACTTACAGTTTGAAC 540  
 DB 1199 GGACATGATACCCCGTGAAGTAACTGCGGCGGCGCTCGTTCACTTACAGTTTGAAC 1252

QY 541 CCGTGGAGGACGGGCGAGCTCGCGGTGCAATGTGTTTACAGCGGTGATGGAGCATGA 600  
DB 1253 CCGTGGAGGACGGGCGAGCTCGCGGTGCAATGTGTTTACAGCGGTGATGGAGCATGA 1312  
QY 601 AGATGCTCGACA CGCTGCGAGAACAGCAGCTAGATTAA CCTAGAAAGATATATATATG 660  
DB 1313 AGATGCTCGACA CGCTGCGAGAACAGCAGCTAGATTAA CCTAGAAAGATATATATATG 1372  
QY 661 TGACGTACGTTAAAGATATATCA TCGGTAAATTTGACGATGGATCC 707  
DB 1373 TGACGTACGTTAAAGATATATCA TCGGTAAATTTGACGATGGATCC 1419

RESULT 7  
ACAS5367  
ID ACAS5367 standard; DNA; 4943 BP.  
ACAS5367;  
XX  
XX  
XX 06-JUN-2003 (first entry)  
XX  
XX  
DE Transformation vector piggyBAC related plasmid PBS-ITR-EGFP.  
KW PiggyBac; transposon; eukaryotic transformation vector; ds;  
KW transformed cell; transformed embryo; transgenic; plasmid; cyclic;  
XX  
XX Synthetic.  
OS  
XX US2002173634-A1.  
PN  
PD 21-NOV-2002.  
XX  
XX 30-OCT-2001; 2001US-00001189.  
PF  
XX  
PR 31-OCT-2000; 2000US-0244677P.  
PR 01-NOV-2000; 2000US-0244984P.  
XX  
XX (FRAS/) FRASER M J.  
PA (LIIX/) Li X.  
PA (BEAM/) BEAM T.  
PA (HUAV/) HUA-VAN A.  
XX  
PI Fraser MJ, Li X, Beam T, Hua-Van A;  
XX WPI; 2003-352597/33.  
XX  
XX  
XX New DNA molecule in the transposon piggyBac, useful for transferring  
PT genes into host cells or embryos for transforming the cells of embryos  
PT that can be used in making transgenic organisms.  
XX  
XX Example 6; Fig 18 (B); 151pp; English.  
PS  
XX  
XX The invention describes a DNA molecule comprising at least 163  
CC consecutive nucleotide base pairs of the 3' terminal region beginning at  
CC the 3' terminal base pair, and at least 125 consecutive nucleotide base  
CC pairs of the 5' terminal region beginning at the 5' terminal base pair, of  
CC the piggyBAC molecule. The region extends from the restriction site SacI  
CC to the end of the piggyBAC molecule. The DNA molecule in the transposon  
CC piggyBac is useful for transferring genes into host cells or embryos for  
CC transforming the cells of embryos. The transformed cells or embryos are  
CC useful for developing or making transgenic organisms. This sequence  
CC represents plasmid used in the creation of minimal sequence eukaryotic  
CC transformation vector piggyBac  
XX  
SQ Sequence 4943 BP; 1221 A; 1305 C; 1245 G; 1172 T; 0 U; 0 Other;

Query Match 96.9%; Score 685; DB 10; Length 4943;  
Best Local Similarity 99.2%; Pred No. 5.9e-206;  
Matches 701; Conservative 0; Mismatches 0; Indels 6; Gaps 1;  
QY 1 GGAATCCATGCGGCAATTTTACGACAGACTATCTTCTAGGGTTAATCACTGATGATCAGG 60

DB 719 GGATCCCATGCGTCAATTTTACGACAGCTATCTTTCTAGGGTTAATCACTGATGATCAGG 778  
QY 61 ATCATATGTTGGGCTTTTTCGGCTCAGTCATCGCCAACTGCGCTATCTGGGCA 120  
DB 779 ATCATATGTTGGGCTTTTTCGGCTCAGTCATCGCCAACTGCGCTATCTGGGCA 838  
QY 121 TCGGGAGGAAGAACCCGCTGCTTTTCCGCGAGGTTGAAGGCGCATGGAAGAGTTTG 180  
DB 839 TCGGGAGGAAGAACCCGCTGCTTTTCCGCGAGGTTGAAGGCGCATGGAAGAGTTTG 898  
QY 181 CCGAGATGACCTGCTGCTGATTTGAGGTGAGGAAAGCAGCTTTACATGATGATTC 240  
DB 899 CCGAGATGACCTGCTGCTGATTTGAGGTGAGGAAAGCAGCTTTACATGATGATTC 958  
QY 241 GGGAAAGGTGCGCCATGCAAGCTTTTAA CGGTGAACCTGTTCTGAGGCACTGGGATTA 300  
DB 959 GGGAAAGGTGCGCCATGCAAGCTTTTAA CGGTGAACCTGTTCTGAGGCACTGGGATTA 1018  
QY 301 CCAAGTTCTGCGCGCTTTTCCGAGACACAGTTCCGATGATGACCCGAAAGCCATCAACA 360  
DB 1019 CCAAGTTCTGCGCGCTTTTCCGAGACACAGTTCCGATGATGACCCGAAAGCCATCAACA 1078  
QY 361 ACCCGAAATATCCGCGCAGCAGCCGAA CTGCGCGTCCGCTGATGATGATGATGATGATG 420  
DB 1079 ACCCGAAATATCCGCGCAGCAGCCGAA CTGCGCGTCCGCTGATGATGATGATGATGATG 1138  
QY 421 GTGCGGCGCTGGGATATTAAGTCAAGCAGGAGACGGGTATCTGCTGATGATGATGATGAT 480  
DB 1139 GTGCGGCGCTGGGATATTAAGTCAAGCAGGAGACGGGTATCTGCTGATGATGATGATGAT 1198  
QY 481 GGACATGATATCCCGTGAAGTTACCCGCGGCGCGCTGCTGATGATGATGATGATGATG 540  
DB 1199 GGACATGATATCCCGTGAAGTTACCCGCGGCGCGCTGCTGATGATGATGATGATGATG 1252  
QY 541 CCGTGGAGGACGGGCGAGCTGCGGTCAATGTGTTTACAGCGATGATGAGCATGA 600  
DB 1253 CCGTGGAGGACGGGCGAGCTGCGGTCAATGTGTTTACAGCGATGATGAGCATGA 1312  
QY 601 AGATGCTCGACACGCTCGAGAACAGCAGCTAGATTAA CCTAGAAAGATATATATATG 660  
DB 1313 AGATGCTCGACACGCTCGAGAACAGCAGCTAGATTAA CCTAGAAAGATATATATATG 1372  
QY 661 TGACGTACGTTAAAGATATATCA TCGGTAAATTTGACGATGGATCC 707  
DB 1373 TGACGTACGTTAAAGATATATCA TCGGTAAATTTGACGATGGATCC 1419

RESULT 8  
ACAS5368  
ID ACAS5368 standard; DNA; 4944 BP.  
ACAS5368;  
XX  
XX  
XX 06-JUN-2003 (first entry)  
XX  
XX  
DE Transformation vector piggyBAC related plasmid PBS-ITR-EGFP.  
KW PiggyBac; transposon; eukaryotic transformation vector; ds;  
KW transformed cell; transformed embryo; transgenic; plasmid; cyclic;  
XX  
XX Synthetic.  
OS  
XX US2002173634-A1.  
PN  
PD 21-NOV-2002.  
XX  
XX 30-OCT-2001; 2001US-00001189.  
PF  
XX  
PR 31-OCT-2000; 2000US-0244677P.  
PR 01-NOV-2000; 2000US-0244984P.  
XX

PA (FRAS/) FRASER M J.  
 PA (LIXX/) LI X.  
 PA (BEAM/) BEAM T.  
 PA (HUA/V/) HUA-VAN A.  
 PI Fraser MJ, Li X, Beam T, Hua-Van A;  
 XX WPI; 2003-352597/33.  
 DR  
 XX  
 PT New DNA molecule in the transposon piggyBac, useful for transferring  
 PT genes into host cells or embryos for transforming the cells of embryos  
 PT that can be used in making transgenic organisms.  
 XX  
 PS Example 6; Fig 19(B); 151pp; English.

CC The invention describes a DNA molecule comprising at least 163  
 CC consecutive nucleotide base pairs of the 3' terminal region beginning at  
 CC the 3' terminal base pair, and at least 125 consecutive nucleotide base  
 CC pairs of the 5' terminal region beginning at the 5' terminal base pair, of  
 CC the piggyBac molecule. The region extends from the restriction site SacI  
 CC to the end of the piggyBac molecule. The DNA molecule in the transposon  
 CC piggyBac is useful for transferring genes into host cells or embryos for  
 CC transforming the cells of embryos. The transformed cells or embryos are  
 CC useful for developing or making transgenic organisms. This sequence  
 CC represents plasmid used in the creation of minimal sequence eukaryotic  
 CC transformation vector piggyBac

XX Sequence 4944 BP; 1221 A; 1302 C; 1247 G; 1174 T; 0 U; 0 Other;

Query Match 96.4%; Score 681.8; DB 10; Length 4944;  
 Best Local Similarity 98.9%; Pred. No. 6.1e-205;

Matches 699; Conservative 0; Mismatches 2; Indels 6; Gaps 1;

QY 1 GGATCCATGCGTCAATTTTACGACAGACTATCTTTGAGGTTATCTAGTCGATCAGG 60  
 DB 719 GGATCCATGCGTCAATTTTACGACAGACTATCTTTGAGGTTATCTAGTCGATCAGG 778  
 QY 61 ATCATATCGTGGGCTTTTCCGGCTCACTCATCGCCCAAGCTGGGCTATCTGGGCA 120  
 DB 779 ATCATATCGTGGGCTTTTCCGGCTCACTCATCGCCCAAGCTGGGCTATCTGGGCA 838  
 QY 121 TCGGGGAGGAAGAACCCGCTCTTTTCCGCGAGTTGAAGCGGATGAAAGATTG 180  
 DB 839 TCGGGGAGGAAGAACCCGCTCTTTTCCGCGAGTTGAAGCGGATGAAAGATTG 898  
 QY 181 CCGAGATGACTGCTGCTGATTCAGTTGACGAAACGACGTTTACATGATGATTC 240  
 DB 899 CCGAGATGACTGCTGCTGATTCAGTTGACGAAACGACGTTTACATGATGATTC 958  
 QY 241 GGGAAAGTGGCCATGACAGCCCTTTTAAAGCTGTAACCTTTGTTGAGGCCACTGGGATA 300  
 DB 959 GGGAAAGTGGCCATGACAGCCCTTTTAAAGCTGTAACCTTTGTTGAGGCCACTGGGATA 1018  
 QY 301 CCACTTGTGCGGCTTTTCCGGAACAGTCCGGAATGTGACGCCGGAAGGCAATCAGCA 360  
 DB 1019 CCACTTGTGCGGCTTTTCCGGAACAGTCCGGAATGTGACGCCGGAAGGCAATCAGCA 1078  
 QY 361 ACCGGAACAATACCGGCGACAGCCGGAATCTGCGGCGATGTGCAATTAATGACAGC 420  
 DB 1079 ACCGGAACAATACCGGCGACAGCCGGAATCTGCGGCGATGTGCAATTAATGACAGC 1138  
 QY 421 GTGGGGGCTGGGATTTAAGTCAAGCAAGAGCGGATCTCTGCTGATGATGCGGCAAT 480  
 DB 1139 GTGGGGGCTGGGATTTAAGTCAAGCAAGAGCGGATCTCTGCTGATGATGCGGCAAT 1198  
 QY 481 GGAATGATATCCCGGATTTAAGTCAAGCAAGAGCGGATCTCTGCTGATGATGCGGCAAT 540  
 DB 1199 GGAATGATATCCCGGATTTAAGTCAAGCAAGAGCGGATCTCTGCTGATGATGCGGCAAT 1252  
 QY 541 CCGTGAAGAGCGGCAAGCTGCGGCTGCAAAATGTTTTCATGAGCGTGAAGAGAGATGA 600  
 DB 1253 CCGTGAAGAGCGGCAAGCTGCGGCTGCAAAATGTTTTCATGAGCGTGAAGAGAGATGA 1312

QY 601 AGATGCTGACAGCGCTGACAGACGAGCTAGATTAACTAGAAAGATATCATATTG 660  
 DB 1313 AGATGCTGACAGCGCTGACAGACGAGCTAGATTAACTAGAAAGATATCATATTG 1372  
 QY 661 TGACGTACGTTAAAGATTAATCATGCTGTAATTAATGACCGATGATCC 707  
 DB 1373 TGACGTACGTTAAAGATTAATCATGCTGTAATTAATGACCGATGATCC 1419

RESULT 9

ADP70043  
 ID ADP70043 standard; DNA; 48502 BP.

AC ADP70043;

DT 26-AUG-2004 (first entry)

DE Nucleic acid detection-related bacteriophage lambda DNA Segid6.

DE Nucleic acid detection; nucleic acid amplification; microchip;

KM Isothermal; chip substrate; db.

OS Bacteriophage lambda.

PN JP2004154008-A.

PD 03-JUN-2004.

PF 01-NOV-2002; 2002JP-00320330.

PR 01-NOV-2002; 2002JP-00320330.

PA (EIKEN ) EIKEN KAGAKU KK.

XX WPI; 2004-445081/42.

DR Detecting nucleic acid such as DNA, cDNA and RNA, involves performing  
 PT nucleic acid amplification in microchip under isothermal conditions, and  
 PT analyzing obtained amplified product on microchip.

PS Example 2; SEQ ID NO 6; 49pp; Japanese.

CC This invention relates to a novel method of detecting nucleic acid which  
 CC comprises performing nucleic acid amplification in a microchip under  
 CC isothermal conditions, and analyzing the obtained amplified product on  
 CC the microchip. The method is useful for detecting nucleic acid and  
 CC enables performing a nucleic acid amplification reaction and its analysis  
 CC continuously on the microchip without affecting the chip substrate. The  
 CC method enables simple, reliable and cost-effective amplification,  
 CC analysis and detection of nucleic acid in a microchip. The invention  
 CC enables detection of trace amounts of nucleic acid with amplification  
 CC efficiency. The present sequence is that of a bacteriophage lambda DNA  
 CC sequence which was used in the exemplification of the invention.

XX Sequence 48502 BP; 12334 A; 11362 C; 12820 G; 11986 T; 0 U; 0 Other;

Query Match 82.8%; Score 585.4; DB 12; Length 48502;  
 Best Local Similarity 99.8%; Pred. No. 7.1e-174;

Matches 586; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 45 ATCTAGCTGATCAGATATCATATCTGCGGCTTTTCCGCTCATCTCCGCCAAGC 104  
 DB 3055 ATCTAGCTGATCAGATATCATATCTGCGGCTTTTCCGCTCATCTCCGCCAAGC 3114  
 QY 105 TGGCGCTATCTGGGATCGGGAGAAAGCCGTCCTTTTCCGCGAGTTGAAGCG 164  
 DB 3115 TGGCGCTATCTGGGATCGGGAGAAAGCCGTCCTTTTCCGCGAGTTGAAGCG 3174  
 QY 165 GCATGAAAGATTGCGCGAGATGACTGCTGCTGATGACGTTGAGCGAAGACGACG 224  
 DB 3175 GCATGAAAGATTGCGCGAGATGACTGCTGCTGATGACGTTGAGCGAAGACGACG 3234  
 QY 225 TTATACATGATGATTCGGAAGAGTGGCCATGACAGCCCTTTTAAAGGTGAATCTGTTG 284



## RESULT 11

AB878924/c

ID AB878924 standard; DNA; 26565 BP.

AC AB878924;

DT 17-DEC-2002 (first entry)

DE E. coli CFT073 genomic sequence #91.

Pathogenic; Escherichia coli CFT073 infection; livestock; pyelonephritis;  
urinary tract infection; open reading frame; ORF; uropathogenic;  
antibacterial; atrophic; nephrotoxic; gene; ds.

OS Escherichia coli.

XX MO200259320-A2.

XX 01-AUG-2002.

XX 19-OCT-2001; 2001WO-US046833.

XX 19-OCT-2000; 2000US-0242412P.

XX (WISC ) WISCONSIN ALUMNI RES FOUND.

XX Blattner FR, Welch RA, Burland VD;

XX WPI; 2002-691532/74.

New DNA sequences of the pathogenic Escherichia coli CFT073 strain,  
useful for preventing or treating E. coli CFT073 infection in humans or  
livestock.

PS Claim 1; Page 276-291; 7655p; English.

The present invention relates to polynucleotide sequences from the genome  
of the pathogenic Escherichia coli strain CFT073. Almost all the  
sequences present in E. coli CFT073 are absent in the previously  
sequenced laboratory strain K-12. The polynucleotide sequences of the  
invention are useful for preventing, diagnosing or treating E. coli  
CFT073 infection in humans or livestock. The polynucleotide sequences are  
useful for preventing urinary tract infections and pyelonephritis.  
Likewise, the polypeptides encoded by the different open reading frames  
(ORF1-5) are useful for generating a vaccine against uropathogenic E.  
coli strains. AB878834-AB879085 represent genomic sequences from E. coli  
strain CFT073

SQ Sequence 26565 BP; 7042 A; 6727 C; 5778 G; 7016 T; 0 U; 2 Other;

Query Match 80.1%; Score 566.2; DB 6; Length 26565;

Best Local Similarity 97.8%; Pred. No. 6.5e-168;

Matches 574; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 45 ATCTAGCTGCATCGAGATCATATGTCGGGTCTTTTTCGCGTCACTGCGCCAGC 104

DB 3221 ATTCAGCTGCATCGAGATCATATGTCGGGTCTTTTTCGCGTCACTGCGCCAGC 3162

QY 105 TGGCGCTATCGGCGATCGGGAGAGAAAGCCGTCCTTTTCCCGAGGTTGAAGCG 164

DB 3161 TGGCGCTATCGGCGATCGGGAGAGAAAGCCGTCCTTTTCCCGAGGTTGAAGCG 3102

QY 165 GCATGGAAGAGTTTGGCGAGATGCTGCTGATTTGACGTTGAGCGAAAGCAGC 224

DB 3101 GCATGGAAGAGTTTGGCGAGATGCTGCTGATTTGACGTTGAGCGAAAGCAGC 3042

QY 225 TTATACCTGATGATTCGGGAAGGTGTGCCATGCAAGCCTTTAAAGGTGAACCTGTTGTT 284

DB 3041 TTCCACCTGATGATTCGGGAAGGTGTGCCATGCAAGCCTTTAAAGGTGAACCTGTTGTT 2982

QY 285 CAGGCAACCTGGATACCAAGTTGTCGGGGCTTTTCCGAGACAGTTCCGAGTGGTACG 344

DB 2981 CAGGCAACCTGGATACCAAGTTGTCGGGGCTTTTCCGAGACAGTTCCGAGTGGTACG 2922

QY 345 CCGAAGCGCATACGAAACCCGAAACATACCGGCGACAGCCGGAACTCCCGTGGGTGTG 404

DB 2921 CCGAAGCGCATACGAAACCCGAAACATACCGGCGACAGCCGGAACTCCCGTGGGTGTG 2862

QY 405 CAGATTATGACAGCGGTGGCGCTGGGATATTACGTCAGCGAGACGGGTATCCCTGGC 464

DB 2861 CAGATTATGACAGCGGTGGCGCTGGGATATTACGTCAGCGAGACGGGTATCCCTGGC 2802

QY 465 TGGATGCCGGAAGAAATGACATGATATCCCGTGAATTACCCGGGCGCGCTCGTTG 524

DB 2801 TGGATGCCGGAAGAAATGACATGATATCCCGTGAATTACCCGGGCGCGCTCGTTG 2742

QY 525 ATTACGTTTTTGAACCCGTGAGAGACGGGCGACACTCGCGGTGAAATGTTTACAGC 584

DB 2741 ATTACGTTTTTGAACCCGTGAGAGACGGGCGACACTCGCGGTGAAATGTTTACAGC 2682

QY 585 GTGATGAGAGATGAAGATGCTCGACAGCTGGAACACGCGACT 631

DB 2681 GTGATGAGAGATGAAGATGCTCGACAGCTGGAACACGCGACT 2635

## RESULT 12

ADH80491/c

ID ADH80491 standard; DNA; 26565 BP.

XX ADH80491;

XX 22-APR-2004 (first entry)

DE Escherichia coli CFT073 genome contig #91.

XX ds; gene; Escherichia coli; CFT073; Escherichia coli CFT073 infection.

XX OS Escherichia coli; CFT073.

XX US2003165870-A1.

XX 04-SEP-2003.

XX 01-MAR-2002; 2002US-00085959.

XX 01-MAR-2002; 2002US-00085959.

XX (BLAT/) BLATTNER F R.

XX (WELC/) WELCH R A.

XX (BURL/) BURLAND V D.

XX Blattner FR, Welch RA, Burland VD;

XX WPI; 2003-863698/80.

XX New nucleic acid of Escherichia coli CFT073, useful for preparing a

XX composition for diagnosing, treating or preventing infection caused by

XX Escherichia coli CFT073.

XX Claim 1; SEQ ID NO 91; 4pp; English.

CC The invention relates to an isolated Escherichia coli CFT073 nucleic acid

CC molecule. The nucleic acid is useful for preparing a composition for

CC diagnosing, treating or preventing infection caused by Escherichia coli

CC CFT073. The present sequence represents a contig of the Escherichia coli

CC CFT073 genome that is not present in Escherichia coli K-12.

SQ Sequence 26565 BP; 7042 A; 6727 C; 5778 G; 7016 T; 0 U; 2 Other;

Query Match 80.1%; Score 566.2; DB 10; Length 26565;

Best Local Similarity 97.8%; Pred. No. 6.5e-168;

Matches 574; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 45 ATCTAGCTGCATCGAGATCATATGTCGGGTCTTTTTCGCGTCACTGCGCCAGC 104

```

Db      3221 ATCAGCTGCATAGATCATATGCTGGGCTCTTTTCCGGCTCAGTCATCGCCCAAGC 3162
Qy      105 TGGCGCTATCTGGGCGATCGGGAGGAAGAACCCCTGCTTTTCCCGGAGGTTGAAGCG 164
Db      3161 TGGCGCTATCTGGGCGATCGGGAGGAAGAACCCCTGCTTTTCCCGGAGGTTGAAGCG 3102
Qy      165 GCATGGAAGAAGTTTGGCGAGATACCTGCTGCTGATGGAAGTTGAGCGAAACGACG 224
Db      3101 GCATGGAAGAAGTTTGGCGAGATACCTGCTGCTGATGGAAGTTGAGCGAAACGACG 3042
Qy      225 TTACCATGATGATTCGGGAAAGTGTGGCGATGACGCTTTAAACGCTGAATCTTGTGTT 284
Db      3041 TTCACCATGATGATTCGGGAAAGTGTGGCGATGATCCTTTAAACGCTGAATCTTGTGTT 2982
Qy      285 CAGGCGACCTGGGATACCAAGTTCCTCGGCTTTTCCGGAACAGATTCGGGATGCTGCTG 344
Db      2981 CAGGCGACCTGGGATACCAAGTTCCTCGGCTTTTCCGGAACAGATTCGGGATGCTGCTG 2922
Qy      345 CCGAAGCGCATCAGCAACCCGAACATATCCGCGACAGCCGGAACCTGCTGCGGTGTG 404
Db      2921 CCGAAGCGCATCAGCAACCCGAACATATCCGCGACAGCCGGAACCTGCTGCGGTGTG 2862
Qy      405 CAGATTATGACAGCGGCTGGGGGCTGGGATTTACGTCAGCGAGGACGGATATCTGTC 464
Db      2861 CAGATTATGACAGCGGCTGGGGGCTGGGATTTACGTCAGCGAGGACGGATATCTGTC 2802
Qy      465 TGGATGCCGCAAGAAATGACATGATGATACCCCGTGAATTAACGTCAGCGAGGACGG 524
Db      2801 TGGATGCCGCAAGAAATGACATGATGATACCCCGTGAATTAACGTCAGCGAGGACGG 2742
Qy      525 ATTACGCTTTTGAACCCGTTGAGAGGACGGGACGACTCGCGTGCMAATGTGTTTACAGC 584
Db      2741 ATTACGCTTTTGAACCCGTTGAGAGGACGGGACGACTCGCGTGCMAATGTGTTTACAGC 582
Qy      585 GTGATGAGACAGATGAAAGTCTCGACACGCTGCGAGAACGACGACGCT 631
Db      2681 GTGATGAGACAGATGAAAGTCTCGACACGCTGCGAGAACGACGACGCT 2635

```

```

RESULT 13
ACD19059
ID      ACD19059 standard; DNA; 46819 BP.
XX
AC      ACD19059;
XX
DT      27-OCT-2003 (revised)
DT      21-AUG-2003 (first entry)
XX
DE      E. coli 0157 unique DNA sequence OZID_72.
XX
KW      OZID; ds; acute haemorrhagic colitis; haemolytic uraemic syndrome;
XX      food poisoning.
XX
OS      Escherichia coli; strain 0157:H7.
XX
PN      US2003023075-A1.
XX
PD      30-JAN-2003.
XX
PF      01-APR-2002; 2002US-00114170.
XX
PR      04-DEC-1998; 98US-0110955P.
PR      03-DEC-1999; 99US-00453702.
XX
PA      (BLAT/) BLATTNER F R.
PA      (BURL/) BURLAND V D.
PA      (PERN/) PERNA N T.
PA      (PLUNK/) PLUNKETT G.
PA      (WELC/) WELCH R.
XX
PI      Blattner FR, Burland VD, Perna NT, Plunkett G, Welch R;

```

```

DR      WPI; 2003-479497/45.
XX
PT      New DNA sequences from Escherichia coli strain 0157:H7, useful for
PT      detecting E. coli 0157:H7 in a sample, or in designing diagnostic probes
PT      which can be used to distinguish strain 0157:H7 from strain K12 using
PT      molecular techniques.
XX
PS      Claim 16; SEQ ID NO 72; 33pp; English.
XX
CC      The invention relates to an isolated DNA molecule comprising an E. coli
CC      strain 0157:H7 sequence selected from a clonal library of a locus of
CC      a urease gene cluster, a RTX toxin-like gene cluster, a locus of
CC      enterocyte effacement and 2 genes from its associated lymphocytic phage
CC      933W (a putative serine/threonine kinase and a tail fibre gene); E. coli
CC      0157:H7 can cause food poisoning, specifically acute haemorrhagic colitis
CC      (which can develop into haemolytic uraemic syndrome). Also included are
CC      an isolated DNA molecule comprising a nucleotide sequence identical to at
CC      least 25 contiguous nucleotides contained in DNA sequences selected from
CC      ACD18988-ACD19242 (being 255 E. coli 0157 DNA sequences which are not
CC      found in E. coli K12), a recombinant DNA construction comprising the DNA
CC      (or distinguishing between 0157 and K12) using a probe derived from one
CC      of the 255 sequences. The DNA sequences are useful in detecting E. coli
CC      0157:H7 in a sample, for the early diagnosis of humans and livestock
CC      infected with 0157:H7, and in designing diagnostic probes which can be
CC      used to distinguish strain 0157:H7 from strain K12 using molecular
CC      techniques. The present sequence is one of the 255 E. coli 0157:H7 DNA
CC      sequence (termed OZID_1-OZID255). Note: The sequence data for this patent
CC      did not form part of the printed specification, but was obtained in
CC      electronic format directly from the USPTO at
CC      segdata.uspto.gov/sequence.html?docid=2003023075 (Updated on 27-OCT-2003
XX      to standardise OS field)
XX
SQ      Sequence 46819 BP; 11858 A; 10855 C; 12774 G; 11317 T; 0 U; 15 Other;

```

Query Match

Beeft Local Similarity 80.1%; Score 566.2; DB 9; Length 46819;

Matches 574; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

```

Qy      45 ATCTAGCTGCATCAGATCATATCGTGGGCTCTTTTCCGCTCAATCATGCCCAAGC 104
Db      14795 ATCCAGCTGCATCAGATCATATCGTGGGCTCTTTTCCGCTCAATCATGCCCAAGC 14854
Qy      105 TGGCGCTATCTGGGCGATCGGGAGGAAGAACCCCTGCTTTTCCCGGAGGTTGAAGCG 164
Db      14855 TGGCGCTATCTGGGCGATCGGGAGGAAGAACCCCTGCTTTTCCCGGAGGTTGAAGCG 14914
Qy      165 GCATGGAAGAAGTTTGGCGAGATACCTGCTGCTGATGGAAGTTGAGCGAAACGACG 224
Db      14915 GCATGGAAGAAGTTTGGCGAGATACCTGCTGCTGATGGAAGTTGAGCGAAACGACG 14974
Qy      225 TTTCACATGATGATTCGGGAAAGTGTGGCGATGACGCTTTTAAACGCTGTAATCTTGTGTT 284
Db      14975 TTTCACATGATGATTCGGGAAAGTGTGGCGATGACGCTTTTAAACGCTGTAATCTTGTGTT 15034
Qy      285 CAGGCGACCTGGGATACCAAGTTCCTCGGCTTTTCCGGAACAGATTCGGGATGCTGCTG 344
Db      15035 CAGGCGACCTGGGATACCAAGTTCCTCGGCTTTTCCGGAACAGATTCGGGATGCTGCTG 15094
Qy      345 CCGAAGCGCATCAGCAACCCGAACATATCCGCGACAGCCGGAACCTGCTGCGGTGTG 404
Db      15095 CCGAAGCGCATCAGCAACCCGAACATATCCGCGACAGCCGGAACCTGCTGCGGTGTG 15154
Qy      405 CAGATTATGACAGCGGCTGGGGGCTGGGATTTACGTCAGCGAGGACGGATATCTGTC 464
Db      15155 CAGATTATGACAGCGGCTGGGGGCTGGGATTTACGTCAGCGAGGACGGATATCTGTC 15214
Qy      465 TGGATGCCGCAAGAAATGACATGATGATACCCCGTGAATTAACGTCAGCGAGGACGG 524
Db      15215 TGGATGCCGCAAGAAATGACATGATGATACCCCGTGAATTAACGTCAGCGAGGACGG 15274
Qy      525 ATTACGCTTTTGAACCCGTTGAGAGGACGGGACGACTCGCGTGCMAATGTGTTTACAGC 584

```



Db 15275 ATTCAGCTCTTGAACCCGAGGAGGAGAGCCGCGGTGCAATGTGTTACAGC 15334  
 QY 585 GTGATGGAGCAGATGAAGATGCTGACACGCTGCGAAGACAGCAGCT 631  
 Db 15335 GTGATGGAGCAGATGAAGATGCTGACACGCTGCGAAGACAGCAGCT 15381

RESULT 14  
 AD00585  
 ID AD00585 standard; DNA; 46897 BP.

AC AD00585;  
 DT 04-DEC-2003 (first entry)  
 XX  
 DE Enterohaemorrhagic E. coli O157:H7-specific nucleic acid SEQ ID NO: 630.  
 XX  
 KW ds; gene; enterohaemorrhagic; anti-bacterial.  
 XX  
 OS Escherichia coli; O157:H7.  
 XX  
 PN JF2002355074-A.  
 XX  
 PD 10-DEC-2002.  
 XX  
 PF 24-JAN-2002; 2002JF-00015959.  
 XX  
 PR 24-JAN-2001; 2001JP-00112010.  
 XX  
 PA (UYTS-) UNIV TSUKUBA.  
 XX  
 DR WPI; 2003-451640/43.

PT Enterohaemorrhagic Escherichia coli O157:H7-specific nucleic acid molecule  
 XX  
 PT and a polypeptide and its use, a polypeptide, a vector and a host cell.  
 XX  
 PS Claim 2; SEQ ID NO 630; 2067bp; Japanese.  
 XX

CC The invention relates to a novel enterohaemorrhagic Escherichia coli  
 CC O157:H7-specific nucleic acid molecule. A polynucleotide of the invention  
 CC has anti-bacterial activity. The polypeptide can be used in detection  
 CC and/or treatment of O157:H7 infection. The nucleotide sequence of the  
 CC genome of Enterohaemorrhagic E. coli O157:H7 was determined. The present  
 CC sequence represents an E. coli O157:H7-specific nucleic acid of the  
 CC invention.  
 CC  
 XX

SQ Sequence 46897 BP; 11872 A; 10948 C; 12731 G; 11346 T; 0 U; 0 Other;

Query Match 79.9%; Score 564.6; DB 10; Length 46897;  
 Best Local Similarity 97.6%; Pred. No. 2.8e-167;  
 Matches 573; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 45 ATCTAGCTGCATCAGATCATATGTCGGGCTCTTTTCCGGCTCAGTCATCGCCCAAGC 104  
 Db 14795 ATCCAGCTGCATCAGATCATATGTCGGGCTCTTTTCCGGCTCAGTCATCGCCCAAGC 14854  
 QY 105 TGGCGCTATCTGGGCTATCGGGAGAGAAACCCGCTCTTTTCCCGGAGCTTGAAACG 164  
 Db 14855 TGGCGCTATCTGGGCTATCGGGAGAGAAACCCGCTCTTTTCCCGGAGCTTGAAACG 14914  
 QY 165 GCATGGAAGAGTTGGCCGAGATGACTGCTGTCATTGACGTTGAGCGAAACGACG 224  
 Db 14915 GCATGGAAGAGTTGGCCGAGATGACTGCTGTCATTGACGTTGAGCGAAACGACG 14974  
 QY 225 TTATCAGTATGATTCGGGAGAGTGCCATGACCGCTTTTAAACGTTGATCTGTCGT 284  
 Db 14975 TTATCAGTATGATTCGGGAGAGTGCCATGACCGCTTTTAAACGTTGATCTGTCGT 15034  
 QY 285 CAGGCCACCTGGGATACCACTGTCGGGGCTTTTCCGGAACAAGTTCCGATGCTGACG 344  
 Db 15035 CAGGCCACCTGGGATACCACTGTCGGGGCTTTTCCGGAACAAGTTCCGATGCTGACG 15094  
 QY 345 CCGAAGCGCATCAGCAACCCGAAACATACCGGCGACGCGGAACTGCGCGGTGTG 404

Db 15095 CCGAAGCGCATCAGCAACCCGAAACATACCGGCGACGCGGAACTGCGCGGTGTG 15154  
 QY 405 CAGATTAATGACACCGGTGCGCGCTGGATATTACGTACAGAGACGGGTATCTGGC 464  
 Db 15155 CAGATTAATGACACCGGTGCGCGCTGGATATTACGTACAGAGACGGGTATCTGGC 15214  
 QY 465 TGGATGCCCGCAAGAAATGACATGATATCCCGTACGTTACCCGGCGGGCGCTCTGTC 524  
 Db 15215 TGGATGCCCGCAAGAAATGACATGATATCCCGTACGTTACCCGGCGGGCGCTCTGTC 15274  
 QY 525 ATTCACGTTTTTGAACCCGTGAGAGACGGGACACTCGCGGTGCAAAATGTTTACAGC 584  
 Db 15275 ATTCACGCTTTTGAACCCGTGAGAGACGGGACACTCGCGGTGCAAAATGTTTACAGC 15334  
 QY 585 GTGATGGAGCAGATGAAGATGCTGACACGCTGCGAAGACGACGCT 631  
 Db 15335 GTGATGGAGCAGATGAAGATGCTGACACGCTGCGAAGACGACGCT 15381

RESULT 15  
 ACD19238/C  
 ID ACD19238 standard; DNA; 22306 BP.  
 XX  
 AC ACD19238;  
 XX  
 DT 27-OCT-2003 (revised)  
 DT 21-AUG-2003 (first entry)  
 XX  
 DE E. coli O157 unique DNA sequence OZ1D\_251.  
 XX  
 KW OZ1D; ds; acute haemorrhagic colitis; haemolytic uraemic syndrome;  
 KW Food poisoning.  
 XX  
 OS Escherichia coli; strain O157:H7.  
 XX  
 PN US2003023075-A1.  
 XX  
 PD 30-JAN-2003.  
 XX  
 PP 01-APR-2002; 2002US-00114170.  
 XX  
 PR 04-DEC-1998; 98US-0110955P.  
 PR 03-DEC-1999; 99US-00453702.  
 XX  
 PA (BLAT/) BLATTNER F R.  
 PA (BURL/) BURLAND V D.  
 PA (PERN/) PERNA N T.  
 PA (PLUN/) PLUNKETT G.  
 PA (WELC/) WELCH R.  
 XX  
 PI Blatter FR, Burland VD, Perna NT, Plunkett G, Welch R;  
 XX  
 XX WPI; 2003-479497/45.

PT New DNA sequences from Escherichia coli strain O157:H7, useful for  
 PT detecting E. coli O157:H7 in a sample, or in designing diagnostic probes  
 PT which can be used to distinguish strain O157:H7 from strain K12 using  
 PT molecular techniques.  
 XX  
 XX

PS Claim 16; SEQ ID NO 251; 33bp; English.

XX The invention relates to an isolated DNA molecule comprising an E. coli  
 CC strain O157:H7 sequence selected from a clonal library of cytotoxin-like gene,  
 CC a urease gene cluster, a RTX toxin-like gene cluster, a locus of  
 CC enterocyte effacement and 2 genes from its associated lymphocytic phage  
 CC 93W (a putative serine/threonine kinase and a tail fibre gene). E. coli  
 CC O157:H7 can cause food poisoning, specifically acute haemorrhagic colitis  
 CC (which can develop into haemolytic uraemic syndrome). Also included are  
 CC an isolated DNA molecule comprising a nucleotide sequence identical to at  
 CC least 25 contiguous nucleotides contained in DNA sequences selected from  
 CC ACD18988-ACD19242 (being 25 E. coli O157 DNA sequences which are not  
 CC found in E. coli K12), a recombinant DNA construction comprising the DNA

above as a method for detecting *E. coli* O157:H7 (ATCC 43895) in a sample or distinguishing between O157 and K12) using a probe derived from one of the 255 sequences. The DNA sequences are useful in detecting *E. coli* O157:H7 in a sample, for the early diagnosis of humans and livestock infected with O157:H7, and in designing diagnostic probes which can be used to distinguish strain O157:H7 from strain K12 using molecular techniques. The present sequence is one of the 255 *E. coli* O157:H7 DNA sequences (termed O21D-1-O21D255). Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from the USPTO at [seqdata.uspto.gov/sequence.html?DocID=20030023075](http://seqdata.uspto.gov/sequence.html?DocID=20030023075) (Updated on 27-OCT-2003 to standardise OS field)

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 26, 2005, 08:26:03 ; Search time 378.34 Seconds  
(without alignments)  
10737.291 Million cell updates/sec

Title: US-10-826-523-40

Perfect score: 707  
Sequence: 1 ggaatccatgcgcgaatttc.....aaatgcacgcagggatcc 707

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4300275 seqs, 2872944193 residues

Total number of hits satisfying chosen parameters: 8600550

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA:  
1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*  
10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:\*  
11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:\*  
12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*  
13: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*  
14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:\*  
15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:\*  
16: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq:\*  
17: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq:\*  
18: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*  
19: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq:\*  
20: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*  
21: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	707	100.0	707	US-10-001-189-40	Sequence 40, App1
2	707	100.0	4613	US-10-001-189-46	Sequence 46, App1
3	707	100.0	8999	US-10-001-189-48	Sequence 48, App1
4	707	100.0	9012	US-10-001-189-49	Sequence 49, App1
5	707	100.0	9013	US-10-001-189-50	Sequence 50, App1
6	685	96.9	4943	US-10-001-189-54	Sequence 54, App1
7	685	96.9	4944	US-10-001-189-55	Sequence 55, App1
8	566.2	80.1	26565	US-10-085-959-91	Sequence 56, App1
9	566.2	80.1	46819	US-10-085-959-91	Sequence 91, App1
10	561.4	79.4	22306	US-10-114-170-72	Sequence 72, App1
11	561.4	79.4	3662	US-10-114-170-251	Sequence 251, App1
12	519.2	73.4	3662	US-10-001-189-41	Sequence 41, App1

C 13	511	72.3	4941	13	US-10-001-189-53	Sequence 53, App1
C 14	511	72.3	4951	13	US-10-001-189-51	Sequence 51, App1
C 15	511	72.3	4952	13	US-10-001-189-52	Sequence 52, App1
C 16	425.8	60.2	1427	18	US-10-363-345A-37517	Sequence 37517, A
C 17	425.8	60.2	1427	18	US-10-363-345A-37518	Sequence 37518, A
C 18	382.6	54.1	1427	18	US-10-363-345A-37519	Sequence 37519, A
C 19	382.6	54.1	1427	18	US-10-363-345A-37520	Sequence 37520, A
C 20	223	31.5	34063	14	US-10-114-170-96	Sequence 96, App1
C 21	222.2	31.4	72480	16	US-10-418-837-2	Sequence 2, App1
C 22	96	13.6	847	10	US-09-798-889-27	Sequence 27, App1
C 23	96	13.6	847	16	US-10-633-680-27	Sequence 27, App1
C 24	78	11.0	9984	13	US-10-001-189-63	Sequence 63, App1
C 25	72.8	10.3	7670	13	US-10-001-189-57	Sequence 57, App1
C 26	40	5.7	43	13	US-10-001-189-19	Sequence 19, App1
C 27	40	5.7	3089	18	US-10-425-115-49726	Sequence 49726, A
C 28	39.8	5.6	676	18	US-10-425-115-91844	Sequence 91844, A
C 29	39.8	5.6	1673	16	US-10-437-963-91699	Sequence 21174, A
C 30	39.8	5.6	1796	16	US-10-425-114-29053	Sequence 29053, A
C 31	39.8	5.6	1872	18	US-10-425-115-91842	Sequence 91842, A
C 32	39	5.5	39	13	US-10-001-189-16	Sequence 16, App1
C 33	38.4	5.4	1473	16	US-10-282-122A-20134	Sequence 20134, A
C 34	36.8	5.2	4187	17	US-10-437-963-91699	Sequence 91699, A
C 35	36.2	5.1	1403	17	US-10-767-701-13724	Sequence 13724, A
C 36	36	5.1	1246	15	US-10-369-493-27114	Sequence 27114, A
C 37	35.4	5.0	2073	11	US-09-758-759-158	Sequence 158, App
C 38	35.4	5.0	96587	11	US-09-997-722-250	Sequence 250, App
C 39	35.4	5.0	109519	11	US-09-758-759-1	Sequence 1, App1
C 40	35.2	5.0	4941	13	US-10-001-189-53	Sequence 53, App1
C 41	35.2	5.0	4951	13	US-10-001-189-51	Sequence 51, App1
C 42	35.2	5.0	4952	13	US-10-001-189-52	Sequence 52, App1
C 43	35	5.0	51	9	US-09-969-852-14	Sequence 14, App1
C 44	34.6	4.9	330	18	US-10-425-115-68281	Sequence 68281, A
C 45	34.6	4.9	887	16	US-10-424-599-40607	Sequence 40607, A

ALIGNMENTS

RESULT 1  
US-10-001-189-40  
Sequence 40, Application US/10001189  
Publication No. US20020173634A1  
GENERAL INFORMATION:  
APPLICANT: FRASER JR., MALCOLM J.  
APPLICANT: LI, XU  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TRANSDUCTION USING  
TITLE OF INVENTION: MINIMAL SEGMENTS OF THE EUKARYOTIC TRANSFORMATION  
FILE REFERENCE: 835910-92098  
CURRENT APPLICATION NUMBER: US/10/001.189  
CURRENT FILING DATE: 2001-10-30  
PRIOR APPLICATION NUMBER: 60/244,984  
PRIOR FILING DATE: 2000-11-01  
PRIOR APPLICATION NUMBER: 60/244,677  
PRIOR FILING DATE: 2000-10-31  
NUMBER OF SEQ ID NOS: 70  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 40  
LENGTH: 707  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: ITR Cartridge  
OTHER INFORMATION: sequence  
US-10-001-189-40  
Query Match 100.0%; Score 707; DB 13; Length 707;  
Best Local Similarity 100.0%; Pred. No. 6,3e-236;  
Matches 707; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
1 GGATCCATGCGCAATTTACGACGACTCTCTTACGAGTTAATCTAGTCATGAGG 60  
|||||

```

Db      1  GGATCCCATGCGTCATTTTACGACATCTTTCTAGGGTTAATCTAGCTGCATCAGG 60
Qy      61  ATCATATCGTGGGCTCTTTTTCGGGCTCAGTCATCGCCAAAGCGGAGCTAATGGGCA 120
Db      61  ATCATATCGTGGGCTCTTTTTCGGGCTCAGTCATCGCCAAAGCTGCATCTAGTGGGCA 120
Qy      121  TCGGGGAGGAAGAACCCGCTGCTTTTCCGCGAGGTTGAAGCGGCATGGAAAGATTG 180
Db      121  TCGGGGAGGAAGAACCCGCTGCTTTTCCGCGAGGTTGAAGCGGCATGGAAAGATTG 180
Qy      181  CCGAGGATGACTGCTGCTGCAATTGACGTTGAGCGAAACGACGTTAACCATGATTC 240
Db      181  CCGAGGATGACTGCTGCTGCAATTGACGTTGAGCGAAACGACGTTAACCATGATTC 240
Qy      241  GGAAGGTGTGGCCATGACAGCCCTTTAACGTTAACTGTTGTTCAAGGCCACTGGGATA 300
Db      241  GGAAGGTGTGGCCATGACAGCCCTTTAACGTTAACTGTTGTTGTTCAAGGCCACTGGGATA 300
Qy      301  CCAATTCGTGCGGGCTTTTTCGGAACAAGTCCGAGTGTGACGCCGGAAGCCATCAGCA 360
Db      301  CCAATTCGTGCGGGCTTTTTCGGAACAAGTCCGAGTGTGACGCCGGAAGCCATCAGCA 360
Qy      361  ACCGGAACAATACCGGAGCAGACCGGAACTGCGGCTGCTGCAATTAATGACAGC 420
Db      361  ACCGGAACAATACCGGAGCAGACCGGAACTGCGGCTGCTGCAATTAATGACAGC 420
Qy      421  GTGCGGCTGGGATTTACGTGACGAGAGAGCGGATCTCTGCTGAGATGCGCGAAT 480
Db      421  GTGCGGCTGGGATTTACGTGACGAGAGAGCGGATCTCTGCTGAGATGCGCGAAT 480
Qy      481  GGAATGATGATACCCCGTAGTTTACCGGCGGCGGCGCTGTTCACTGACGTTTGAAC 540
Db      481  GGAATGATGATACCCCGTAGTTTACCGGCGGCGGCGCTGTTCACTGACGTTTGAAC 540
Qy      541  CCGTGAAGAGCGGAGACTCGGAGTAAATGTGTTTACAGGTTGAGAGAGATGA 600
Db      541  CCGTGAAGAGCGGAGACTCGGAGTAAATGTGTTTACAGGTTGAGAGAGATGA 600
Qy      601  AGATGCTGACAGCGCTGAGAACAGCGAGCTAGATTAACTTAGAAAGATATCATATTG 660
Db      601  AGATGCTGACAGCGCTGAGAACAGCGAGCTAGATTAACTTAGAAAGATATCATATTG 660
Qy      661  TGAAGTAAAGATATCATATGCTGTTAAATGACGATGGATCC 707
Db      661  TGAAGTAAAGATATCATATGCTGTTAAATGACGATGGATCC 707

```

## RESULT 2

```

/ Sequence 46, Application US/10001189
/ Publication No. US20020173634A1
/ GENERAL INFORMATION:
/ APPLICANT: FRASER JR., MALCOLM J.
/ APPLICANT: LI, XU
/ TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TRANSPOSITION USING
/ TITLE OF INVENTION: MINIMAL SEGMENTS OF THE EUKARYOTIC TRANSFORMATION
/ FILE REFERENCE: 835910-92098
/ CURRENT APPLICATION NUMBER: US/10/001,189
/ PRIOR FILING DATE: 2001-10-30
/ PRIOR APPLICATION NUMBER: 60/244,984
/ PRIOR FILING DATE: 2000-11-01
/ PRIOR APPLICATION NUMBER: 60/244,677
/ NUMBER OF SEQ ID NOS: 70
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 46
/ LENGTH: 4613
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence: pCRIT-ITR

```

```

/ OTHER INFORMATION: sequence
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (344)..(922)
US-10-001-189-46

```

```

Query Match      100.0%; Score 707; DB 13; Length 4613;
Best Local Similarity 100.0%; Pred. No. 1,7e-235;
Matches 707; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1  GGATCCCATGCGTCATTTTACGACATCTTTCTAGGGTTAATCTAGCTGCATCAGG 60
Db      294  GGATCCCATGCGTCATTTTACGACATCTTTCTAGGGTTAATCTAGCTGCATCAGG 353
Qy      61  ATCATATCGTGGGCTCTTTTTCGGGCTCAGTCATCGCCAAAGCGGAGCTAATGGGCA 120
Db      354  ATCATATCGTGGGCTCTTTTTCGGGCTCAGTCATCGCCAAAGCGGAGCTAATGGGCA 413
Qy      121  TCGGGGAGGAAGAACCCGCTGCTTTTCCGCGAGGTTGAAGCGGCATGGAAAGATTG 180
Db      414  TCGGGGAGGAAGAACCCGCTGCTTTTCCGCGAGGTTGAAGCGGCATGGAAAGATTG 473
Qy      181  CCGAGGATGACTGCTGCTGCAATTGACGTTGAGCGAAACGACGTTAACCATGATTC 240
Db      474  CCGAGGATGACTGCTGCTGCAATTGACGTTGAGCGAAACGACGTTAACCATGATTC 533
Qy      241  GGAAGGTGTGGCCATGACAGCCCTTTAACGTTAACTGTTGTTGTTCAAGGCCACTGGGATA 300
Db      534  GGAAGGTGTGGCCATGACAGCCCTTTAACGTTAACTGTTGTTGTTCAAGGCCACTGGGATA 593
Qy      301  CCAATTCGTGCGGGCTTTTTCGGAACAAGTCCGAGTGTGACGCCGGAAGCCATCAGCA 360
Db      594  CCAATTCGTGCGGGCTTTTTCGGAACAAGTCCGAGTGTGACGCCGGAAGCCATCAGCA 653
Qy      361  ACCGGAACAATACCGGAGCAGACCGGAACTGCGGCTGCTGCAATTAATGACAGC 420
Db      654  ACCGGAACAATACCGGAGCAGACCGGAACTGCGGCTGCTGCAATTAATGACAGC 713
Qy      421  GTGCGGCTGGGATTTACGTGACGAGAGAGCGGATCTCTGCTGAGATGCGCGAAT 480
Db      714  GTGCGGCTGGGATTTACGTGACGAGAGAGCGGATCTCTGCTGAGATGCGCGAAT 773
Qy      481  GGAATGATGATACCCCGTAGTTTACCGGCGGCGGCGCTGTTCACTGACGTTTGAAC 540
Db      774  GGAATGATGATACCCCGTAGTTTACCGGCGGCGGCGCTGTTCACTGACGTTTGAAC 833
Qy      541  CCGTGAAGAGCGGAGACTCGGAGTAAATGTGTTTACAGGTTGAGAGAGATGA 600
Db      834  CCGTGAAGAGCGGAGACTCGGAGTAAATGTGTTTACAGGTTGAGAGAGATGA 893
Qy      601  AGATGCTGACAGCGCTGAGAACAGCGAGCTAGATTAACTTAGAAAGATATCATATTG 660
Db      894  AGATGCTGACAGCGCTGAGAACAGCGAGCTAGATTAACTTAGAAAGATATCATATTG 953
Qy      661  TGAAGTAAAGATATCATATGCTGTTAAATGACGATGGATCC 707
Db      954  TGAAGTAAAGATATCATATGCTGTTAAATGACGATGGATCC 1000

```

## RESULT 3

```

/ Sequence 48, Application US/10001189
/ Publication No. US20020173634A1
/ GENERAL INFORMATION:
/ APPLICANT: FRASER JR., MALCOLM J.
/ APPLICANT: LI, XU
/ TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TRANSPOSITION USING
/ TITLE OF INVENTION: MINIMAL SEGMENTS OF THE EUKARYOTIC TRANSFORMATION
/ FILE REFERENCE: 835910-92098
/ CURRENT APPLICATION NUMBER: US/10/001,189
/ CURRENT FILING DATE: 2001-10-30

```

PRIOR APPLICATION NUMBER: 60/244,984  
 PRIOR FILING DATE: 2000-11-01  
 PRIOR APPLICATION NUMBER: 60/244,677  
 PRIOR FILING DATE: 2000-10-31  
 NUMBER OF SEQ ID NOS: 70  
 SOFTWARE: PatentIn Ver. 2.1  
 SEQ ID NO 48  
 LENGTH: 8999  
 TYPE: DNA  
 ORGANISM: Artificial Sequence  
 FEATURE:  
 OTHER INFORMATION: Description of Artificial Sequence: p(PZ)-Bac-EXFP  
 OTHER INFORMATION: sequence  
 US-10-001-189-48

Query Match 100.0%; Score 707; DB 13; Length 8999;  
 Best Local Similarity 100.0%; Pred. No. 2,4e-235;  
 Matches 707; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGAATCCATGCGTCAATTTTACGACATCTTTCTTACGGTTTATCTAGCTGATCAGG 60  
 DB 8950 GGATCCCATGCGTCAATTTTACGACATCTTTCTTACGGTTTATCTAGCTGATCAGG 8891  
 QY 61 ATCATATGTCGGGCTTTTTCGGGCTCAGTCAATCCGCAAGCTGGGCTATCTGGGCA 120  
 DB 8890 ATCATATGTCGGGCTTTTTCGGGCTCAGTCAATCCGCAAGCTGGGCTATCTGGGCA 8831  
 QY 121 TCGGGGAGAGAAAGACCCGCTCTTTTCCCGAGGTTGAAGCGGCATGGAAGAGTTTG 180  
 DB 8830 TCGGGGAGAGAAAGACCCGCTCTTTTCCCGAGGTTGAAGCGGCATGGAAGAGTTTG 8771  
 QY 181 CCGAGATGATGCTGCTGCAATTTGAGTGAAGCAAGCAAGTTTACATGATGATTC 240  
 DB 8770 CCGAGATGATGCTGCTGCAATTTGAGTGAAGCAAGCAAGTTTACATGATGATTC 8711  
 QY 241 GGGAAAGGTGGGCGCATGACGCTTTTAAAGGTGAAGCTTTGTTCAAGGCCACTGGGATA 300  
 DB 8710 GGGAAAGGTGGGCGCATGACGCTTTTAAAGGTGAAGCTTTGTTCAAGGCCACTGGGATA 8651  
 QY 301 CCAAGTGTGCGGCTTTTCCGGAACAAGTTCGGATGTCAGCCGCAAGCGCATCAGCA 360  
 DB 8650 CCAAGTGTGCGGCTTTTCCGGAACAAGTTCGGATGTCAGCCGCAAGCGCATCAGCA 8591  
 QY 361 ACCCGAACAATACCGGCGGACAGCTGCGGCTGCTGCAATTTATGACAGCG 420  
 DB 8590 ACCCGAACAATACCGGCGGACAGCTGCGGCTGCTGCAATTTATGACAGCG 8531  
 QY 421 GTGCGGGCTGGGATTTAAGTTCAGCGAGGAGCGGGTATCTGGCTGATGCGGAGAAAT 480  
 DB 8530 GTGCGGGCTGGGATTTAAGTTCAGCGAGGAGCGGGTATCTGGCTGATGCGGAGAAAT 8471  
 QY 481 GGAATGGAATACCCCGTGAATTAACCGGCGGCGGCTCGTTCAATTCAGCTTTTGAAC 540  
 DB 8470 GGAATGGAATACCCCGTGAATTAACCGGCGGCGGCTCGTTCAATTCAGCTTTTGAAC 8411  
 QY 541 CCGTGAAGAGCGGCGGACAGCTGCGGCTGCAAAATGTTTTCACAGCTGATGAGCAGATGA 600  
 DB 8410 CCGTGAAGAGCGGCGGACAGCTGCGGCTGCAAAATGTTTTCACAGCTGATGAGCAGATGA 8351  
 QY 601 AGATGCTGACAGCGCTGACAGACAGCGAGTAACTTAAGCAAGTAAATATATATTG 660  
 DB 8350 AGATGCTGACAGCGCTGACAGACAGCGAGTAACTTAAGCAAGTAAATATATATTG 8291  
 QY 661 TGAAGTACGTTAAAGATATCATCGTAAATTTGACCGATGGATCC 707  
 DB 8290 TGAAGTACGTTAAAGATATCATCGTAAATTTGACCGATGGATCC 8244

RESULT 4  
 US-10-001-189-49/c  
 Sequence 49, Application US/10001189  
 Publication No. US20020173634A1  
 GENERAL INFORMATION:

APPLICANT: FRASER JR., MALCOLM J.  
 APPLICANT: LI, XU  
 TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TRANSPOSITION USING  
 TITLE OF INVENTION: MINIMAL SEGMENTS OF THE EUKARYOTIC TRANSFORMATION  
 TITLE OF INVENTION: VECTOR PIGGYBAC  
 FILE REFERENCE: 835910-92098  
 CURRENT APPLICATION NUMBER: US/10/001,189  
 PRIOR FILING DATE: 2001-10-30  
 PRIOR APPLICATION NUMBER: 60/244,984  
 PRIOR FILING DATE: 2000-11-01  
 PRIOR APPLICATION NUMBER: 60/244,677  
 PRIOR FILING DATE: 2000-10-31  
 NUMBER OF SEQ ID NOS: 70  
 SOFTWARE: PatentIn Ver. 2.1  
 SEQ ID NO 49  
 LENGTH: 9012  
 TYPE: DNA  
 ORGANISM: Artificial Sequence  
 FEATURE:  
 OTHER INFORMATION: Description of Artificial Sequence: p(PZ)-Bac-ECFP  
 OTHER INFORMATION: sequence  
 US-10-001-189-49

Query Match 100.0%; Score 707; DB 13; Length 9012;  
 Best Local Similarity 100.0%; Pred. No. 2,4e-235;  
 Matches 707; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGAATCCATGCGTCAATTTTACGACATCTTTCTTACGGTTTATCTAGCTGATCAGG 60  
 DB 8963 GGATCCCATGCGTCAATTTTACGACATCTTTCTTACGGTTTATCTAGCTGATCAGG 8904  
 QY 61 ATCATATGTCGGGCTTTTTCGGGCTCAGTCAATCCGCAAGCTGGGCTATCTGGGCA 120  
 DB 8903 ATCATATGTCGGGCTTTTTCGGGCTCAGTCAATCCGCAAGCTGGGCTATCTGGGCA 8844  
 QY 121 TCGGGGAGAGAAAGACCCGCTCTTTTCCCGAGGTTGAAGCGGCATGGAAGAGTTTG 180  
 DB 8843 TCGGGGAGAGAAAGACCCGCTCTTTTCCCGAGGTTGAAGCGGCATGGAAGAGTTTG 8784  
 QY 181 CCGAGATGATGCTGCTGCAATTTGAGTGAAGCAAGCAAGTTTACATGATGATTC 240  
 DB 8783 CCGAGATGATGCTGCTGCAATTTGAGTGAAGCAAGCAAGTTTACATGATGATTC 8724  
 QY 241 GGGAAAGGTGGGCGCATGACGCTTTTAAAGGTGAAGCTTTGTTCAAGGCCACTGGGATA 300  
 DB 8723 GGGAAAGGTGGGCGCATGACGCTTTTAAAGGTGAAGCTTTGTTCAAGGCCACTGGGATA 8664  
 QY 301 CCAAGTGTGCGGCTTTTCCGGAACAAGTTCGGATGTCAGCCGCAAGCGCATCAGCA 360  
 DB 8663 CCAAGTGTGCGGCTTTTCCGGAACAAGTTCGGATGTCAGCCGCAAGCGCATCAGCA 8604  
 QY 361 ACCCGAACAATACCGGCGGACAGCTGCGGCTGCTGCAATTTATGACAGCG 420  
 DB 8603 ACCCGAACAATACCGGCGGACAGCTGCGGCTGCTGCAATTTATGACAGCG 8544  
 QY 421 GTGCGGGCTGGGATTTAAGTTCAGCGAGGAGCGGGTATCTGGCTGATGCGGAGAAAT 480  
 DB 8543 GTGCGGGCTGGGATTTAAGTTCAGCGAGGAGCGGGTATCTGGCTGATGCGGAGAAAT 8484  
 QY 481 GGAATGGAATACCCCGTGAATTAACCGGCGGCGGCTCGTTCAATTCAGCTTTTGAAC 540  
 DB 8483 GGAATGGAATACCCCGTGAATTAACCGGCGGCGGCTCGTTCAATTCAGCTTTTGAAC 8424  
 QY 541 CCGTGAAGAGCGGCGGACAGCTGCGGCTGCAAAATGTTTTCACAGCTGATGAGCAGATGA 600  
 DB 8423 CCGTGAAGAGCGGCGGACAGCTGCGGCTGCAAAATGTTTTCACAGCTGATGAGCAGATGA 8364  
 QY 601 AGATGCTGACAGCGCTGACAGACAGCGAGTAACTTAAGCAAGTAAATATATATTG 660  
 DB 8363 AGATGCTGACAGCGCTGACAGACAGCGAGTAACTTAAGCAAGTAAATATATATTG 8304  
 QY 661 TGAAGTACGTTAAAGATATCATCGTAAATTTGACCGATGGATCC 707

Db 8303 TGACGTACGTTAAAGATATCATGCTAAATTAATGACCATGGATCC 8257

RESULT 5  
US-10-001-189-50/c  
Sequence 50, Application US/10001189  
Publication No. US20020173634A1  
GENERAL INFORMATION:  
APPLICANT: FRASER JR., MALCOLM J.  
APPLICANT: LI, XU  
APPLICANT: BEAM, TERESA  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TRANSPOSITION USING  
TITLE OF INVENTION: MINIMAL SEGMENTS OF THE EUKARYOTIC TRANSFORMATION  
FILE REFERENCE: 835910-92098  
CURRENT APPLICATION NUMBER: US/10/001,189  
PRIOR FILING DATE: 2001-10-30  
PRIOR FILING DATE: 2000-11-01  
PRIOR FILING DATE: 2000-11-01  
PRIOR FILING DATE: 2000-11-01  
NUMBER OF SEQ ID NOS: 70  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 50  
LENGTH: 9013  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: p(P2)-Bac-EGFP  
US-10-001-189-50

Query Match  
Best Local Similarity 100.0%; Score 707; DB 13; Length 9013;  
Matches 707; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGATCCCATGCTCAATTTTACGACAGATCTTTCTAGGGTTAATCACTGATAGG 60  
Db 8964 GGATCCCATGCTCAATTTTACGACAGATCTTTCTAGGGTTAATCACTGATAGG 60  
Qy 61 ATCATATCGTCGGGCTTTTTCGGGCTGATCGCCCAAGCTGCGCTATCTGGGA 120  
Db 8904 ATCATATCGTCGGGCTTTTTCGGGCTGATCGCCCAAGCTGCGCTATCTGGGA 120  
Qy 121 TCGGGAGGAGAAACCCGTCCTTTTCCCGAGTTGAAGGGGATGGAAGAATTG 180  
Db 8844 TCGGGAGGAGAAACCCGTCCTTTTCCCGAGTTGAAGGGGATGGAAGAATTG 180  
Qy 181 CCGAGATGACTGCTGCTGATGACGTGAGGAAAGCAAGTTTACATGATGATTC 240  
Db 8784 CCGAGATGACTGCTGCTGATGACGTGAGGAAAGCAAGTTTACATGATGATTC 240  
Qy 241 GGGAAAGTGTGCGATGACGCTTTTACGTTAAGCTGTTGCTGAGGCACTGGGATA 300  
Db 8724 GGGAAAGTGTGCGATGACGCTTTTACGTTAAGCTGTTGCTGAGGCACTGGGATA 300  
Qy 301 CCAATTGTGTCGGGCTTTTCCGAGACAGTTCCGATGATGATGATGATGATG 360  
Db 8664 CCAATTGTGTCGGGCTTTTCCGAGACAGTTCCGATGATGATGATGATGATG 360  
Qy 361 ACCGGAACATATCCGCGACAGCCGGAATGCTGCGGTGATGATGATGATGATG 420  
Db 8604 ACCGGAACATATCCGCGACAGCCGGAATGCTGCGGTGATGATGATGATGATG 420  
Qy 421 GTGCGGCGCTGGATATTAATGTCAGCGAGGACGGGATCTGCTGATGATGATG 480  
Db 8544 GTGCGGCGCTGGATATTAATGTCAGCGAGGACGGGATCTGCTGATGATGATG 480  
Qy 481 GGACATGATATCCCGGTGATGATGATGATGATGATGATGATGATGATGATG 540  
Db 8484 GGACATGATATCCCGGTGATGATGATGATGATGATGATGATGATGATGATG 540

Qy 541 CCGTGAAGGACGGGACAGCTCGCGTGGCAATGCTTTTACAGCTGATGAGACAGATGA 600  
Db 8424 CCGTGAAGGACGGGACAGCTCGCGTGGCAATGCTTTTACAGCTGATGAGACAGATGA 600  
Qy 601 AGATCTGACACCGCTGAGAACCGGACGCTAGATTAACCTTGAAGAATTAATGATG 660  
Db 8364 AGATCTGACACCGCTGAGAACCGGACGCTAGATTAACCTTGAAGAATTAATGATG 660  
Qy 661 TGACGTACGTTAAAGATATCATGCTAAATTAATGACCATGGATCC 707  
Db 8304 TGACGTACGTTAAAGATATCATGCTAAATTAATGACCATGGATCC 8258

RESULT 6  
US-10-001-189-54  
Sequence 54, Application US/10001189  
Publication No. US20020173634A1  
GENERAL INFORMATION:  
APPLICANT: FRASER JR., MALCOLM J.  
APPLICANT: LI, XU  
APPLICANT: BEAM, TERESA  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TRANSPOSITION USING  
TITLE OF INVENTION: MINIMAL SEGMENTS OF THE EUKARYOTIC TRANSFORMATION  
FILE REFERENCE: 835910-92098  
CURRENT APPLICATION NUMBER: US/10/001,189  
PRIOR FILING DATE: 2001-10-30  
PRIOR FILING DATE: 2000-11-01  
PRIOR FILING DATE: 2000-11-01  
PRIOR FILING DATE: 2000-11-01  
NUMBER OF SEQ ID NOS: 70  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 54  
LENGTH: 4943  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: pB8-JTR-BCFP  
US-10-001-189-54

Query Match  
Best Local Similarity 96.9%; Score 685; DB 13; Length 4943;  
Matches 701; Conservative 0; Mismatches 0; Indels 6; Gaps 1;

Qy 1 GGATCCCATGCTCAATTTTACGACAGATCTTTCTAGGGTTAATCACTGATAGG 60  
Db 719 GGATCCCATGCTCAATTTTACGACAGATCTTTCTAGGGTTAATCACTGATAGG 60  
Qy 61 ATCATATCGTCGGGCTTTTTCGGGCTGATCGCCCAAGCTGCGCTATCTGGGA 120  
Db 779 ATCATATCGTCGGGCTTTTTCGGGCTGATCGCCCAAGCTGCGCTATCTGGGA 120  
Qy 121 TCGGGAGGAGAAACCCGTCCTTTTCCCGAGTTGAAGGGGATGGAAGAATTG 180  
Db 839 TCGGGAGGAGAAACCCGTCCTTTTCCCGAGTTGAAGGGGATGGAAGAATTG 180  
Qy 181 CCGAGATGACTGCTGCTGATGACGTGAGGAAAGCAAGTTTACATGATGATTC 240  
Db 899 CCGAGATGACTGCTGCTGATGACGTGAGGAAAGCAAGTTTACATGATGATTC 240  
Qy 241 GGGAAAGTGTGCGATGACGCTTTTACGTTAAGCTGTTGCTGAGGCACTGGGATA 300  
Db 959 GGGAAAGTGTGCGATGACGCTTTTACGTTAAGCTGTTGCTGAGGCACTGGGATA 300  
Qy 301 CCAATTGTGTCGGGCTTTTCCGAGACAGTTCCGATGATGATGATGATGATGATG 360  
Db 1019 CCAATTGTGTCGGGCTTTTCCGAGACAGTTCCGATGATGATGATGATGATGATG 360  
Qy 361 ACCGGAACATATCCGCGACAGCCGGAATGCTGCGGTGATGATGATGATGATGATG 420  
Db 1079 ACCGGAACATATCCGCGACAGCCGGAATGCTGCGGTGATGATGATGATGATGATG 420



Qy	421	GTGGGGCGCTGGGATATTAACGTACAGCGAGACGGGTATCTGGCTGTATGCGCGAGAAAT	480
Db	1139	GTGGGGCGCTGGGATATTAACGTACAGCGAGACGGGTATCTGGCTGTATGCGCGAGAAAT	1198
Qy	481	GGACATGAGTACCCCGGTAGTTACCCGGCGGGCGGCGCTCGTTCATTCACGTTTTTGTAAC	540
Db	1199	GGACATGAGTACCCCGGTAGTTACCCGGCGG-----CTCGTTCATTCACGTTTTTGAAAC	1252
Qy	541	CCGTGGAGGACGGGCGAGCTCGCGGTGGAATTCGTTTTTAACAGCTGATGAGACGAGATGA	600
Db	1253	CCGTGGAGGACGGGCGAGCTCGCGGTGGAATTCGTTTTTAACAGCTGATGAGACGAGATGA	1312
Qy	601	AGATGCTCGACACGCTGCAGAACACGACGTAGATTAACCTTAGAAAGATATCATATTG	660
Db	1313	AGATGCTCGACACGCTGCAGAACACGACGTAGATTAACCTTAGAAAGATATCATATTG	1372
Qy	661	TGACGTACGTTAAAGATATATCATGCGTAAATTTGACGATGGGATCC	707
Db	1373	TGACGTACGTTAAAGATATATCATGCGTAAATTTGACGATGGGATCC	1419

## RESULT 7

US-10-001-189-55  
Sequence 55, Application US/10001189  
Publication No. US20020173634A1  
GENERAL INFORMATION:  
APPLICANT: FRASER JR., MALCOLM J.  
APPLICANT: LI, XU  
APPLICANT: BEAM, TERESA  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TRANSCRIPTION USING  
TITLE OF INVENTION: MINIMAL SEGMENTS OF THE EUKARYOTIC TRANSCRIPTION  
FILE REFERENCE: 835910-92098  
CURRENT APPLICATION NUMBER: US/10/001,189  
CURRENT FILING DATE: 2001-10-30  
PRIOR APPLICATION NUMBER: 60/244,984  
PRIOR FILING DATE: 2000-11-01  
PRIOR APPLICATION NUMBER: 60/244,677  
PRIOR FILING DATE: 2000-10-31  
NUMBER OF SEQ ID NOS: 70  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 55  
LENGTH: 4944  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: PBS-ITR-EGFP  
US-10-001-189-55

Query Match	96.9%	Score 685	DB 13	Length 4944
Best Local Similarity	99.2%	Pred. No. 8.6e-228		
Matches 701	Conservative	0	Mismatches 0	Indels 6
				Gaps 1

OY	1	GGATCCCATGGGTCAATTTTACGAGCATATCTTTCTAGAGGTATATCTAGCTGCATCAG	60
Db	719	GGATCCCATGGGTCAATTTTACGAGCATATCTTTCTAGAGGTATATCTAGCTGCATCAG	778
OY	61	ATCATATCTGCGGGCTCTTTTTCCGGCTCAGTCATGCGCCCAAGTCGGCTATCTGGGCA	120
Db	779	ATCATATGTCGGGGCTTTTTTCCGGCTCAGTCATGCGCCCAAGTCGGGGCTATCTGGGCA	838
OY	121	TCGGGGAGAGAAAGCCCGTCCTTTTCCCGGAGTTGAAGCGGCATGAAAGAGATTG	180
Db	839	TCGGGGAGAGAAAGCCCGTCCTTTTCCCGGAGTTGAAGCGGCATGAAAGAGATTG	898
OY	181	CCGAGGAGACATCGCTGTCGACATGACGTTGAGCGGAAAACGACGCTTTACCATGATGATTC	240
Db	899	CCGAGGAGACATCGCTGTCGACATGACGTTGAGCGGAAAACGACGCTTTACCATGATGATTC	958
OY	241	GGGAAAGTGTGGCCATGCACGCGCTTTAAACGTGAACGTTCGTTCAAGGCCACTTGGGATA	300

Db	959	GGAAAGGTGTGGCCATCGACGCTTTAAACGTAAACGTTCGTTACAGCCACCTGGGATA	1018
Qy	301	CCAGTCTGTGCGGGCTTTTCCGGAACACAGTTCGCGATGTCAGGCCGAAACGATCAGCA	360
Db	1019	CCAATTGGTGGCGGCTTTTCCGGAACAGTTTCCGGATGTCAGCCGGAACGATCAGCA	1078
Qy	361	ACCCGAACAATACCGGCGACAGCCGGAACTGCGTGCCTGGTGTGCAGATTAAATGACAGCG	420
Db	1079	ACCCGAACAATACCGGCGACAGCCGGAACTGCGTGCCTGGTGTGCAGATTAAATGACAGCG	1138
Qy	421	GTGCGGGGCTGGGATATTACGTACAGCGAAGACGGGTATCTCGGTGGATGCCGACAAAT	480
Db	1139	GTGCGGGGCTGGGATATTACGTACAGCGAAGACGGGTATCTCGGTGGATGCCGACAAAT	1198
Qy	481	GGAACATGGATACCCCGTGAATTACCCGGCGGCGCGCTTGTCATTCAGTTTTGAAC	540
Db	1199	GGAACATGGATACCCCGTGAATTACCCGGCGG-----CTGTTTCATTCAAGTTTTGAAC	1252
Qy	541	CCGTGGAGGACCGGACACACTCGCGGTGCAAATGTGTTTACACGTGTATGAGCGATGA	600
Db	1253	CCGTGGAGGACCGGACACACTCGCGGTGCAAATGTGTTTACACGTGTATGAGCGATGA	1312
Qy	601	AGATGCTCGACAGCGTCAGAAACGAGCTAGATTAACTCTGAAGATTAATCATTTG	660
Db	1313	AGATGCTCGACAGCGTCAGAAACGAGCTAGATTAACTCTGAAGATTAATCATTTG	1372
Qy	661	TGACGTAACGTTAAAGATTAATCATGCGTAAATTAAGCATGGGATCC	707
Db	1373	TGACGTAACGTTAAAGATTAATCATGCGTAAATTAAGCATGGGATCC	1419

## RESULT 8

```

US-10-001-189-56
: Sequence 56, Application US/10001189
: Publication No. US20020173534A1
: GENERAL INFORMATION:
: APPLICANT: FRASER JR., MALCOLM J.
: APPLICANT: LI, XU
: APPLICANT: BEAM, TERESA
: TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TRANSPOSITION USING
: TITLE OF INVENTION: MINIMAL SEGMENTS OF THE EUKARYOTIC TRANSFORMATION
: TITLE OF INVENTION: VECTOR PIGG1BAC
: FILE REFERENCE: 835910-92098
: CURRENT APPLICATION NUMBER: US/10/001,189
: CURRENT FILING DATE: 2001-10-30
: PRIOR APPLICATION NUMBER: 60/244,984
: PRIOR FILING DATE: 2000-11-01
: PRIOR APPLICATION NUMBER: 60/244,677
: PRIOR FILING DATE: 2000-10-31
: NUMBER OF SEQ ID NOS: 70
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 56
: LENGTH: 4944
: TYPE: DNA
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: Description of Artificial Sequence: pBS-ITR-EYFP
: OTHER INFORMATION: sequence
US-10-001-189-56

```

Query Match	96.9%	Score 685	DB 13	Length 4944
Best Local Similarity	99.2%	Pred. No. 8	6e-226	
Matches 701	Conservative	0	Mismatches	0
			Indels	6
			Gaps	1

OY	GGATCCCATGCGTCAATTTTATCGAGACATCTTTCTAGGGTTAATCATGCGGATCAGG	60
Db	719 GGATCCCATGCGTCAATTTTATCGAGACATCTTTCTAGGGTTAATCATGCGGATCAGG	778
OY	61 ATCATATCGTGGGCTCTTTTTCGGGCTCAGTCATCGCCCAAGCTGGCGCTATCTGGGCA	120
Db	779 ATCATATCGTGGGCTCTTTTTCGGGCTCAGTCATCGCCCAAGCTGGCGCTATCTGGGCA	838
OY	121 TCGGGAGAGAAAGCCCGCTCTTTTCCCGCGAGGTGTGAAGCGGCATGTGAAGAAGTTTG	180

```

1 // Sequence 91, Application US/10085959
2 // Publication No. US20030165870A1
3 // GENERAL INFORMATION:
4 // APPLICANT: Blattner, Frederick R.
5 // APPLICANT: Welch, Rodney A.
6 // APPLICANT: Burland, Valerie D.
7 // TITLE OF INVENTION: No. US20030165870A1el Sequence of E. coli CFT073
8 // FILE REFERENCE: 960296, 97648
9 // CURRENT APPLICATION NUMBER: US/10/085,959
10 // CURRENT FILING DATE: 2002-03-01
11 // PRIOR APPLICATION NUMBER: 60/242,412
12 // PRIOR FILING DATE: 2000-01-19
13 // NUMBER OF SEQ ID NOS: 255
14 // SOFTWARE: Patentin version 3.1
15 // SEQ ID NO 91
16 // LENGTH: 26565
17 // TYPE: DNA
18 // ORGANISM: Escherichia coli
19 // FEATURES:
20 // NAME/KEY: misc_feature
21 // LOCATION: (14853)..(14853)
22 // OTHER INFORMATION: Unsure
23 // IS-10-085-959-91

```

QY	45	ATCTAGCTGCATCAGGATCATATAGTGGGCTCTTTTTCGGAGCTAGTCATCGCCCAAC	104
Db	3221	ATCCAGTGCATCAGGATCATATAGTGGGCTCTTTTTCGGAGCTAGTCATCGCCCAAC	3163

QY	105	TGGCGCATCTGGGCAATCCGGGAGAGAAAGACCCTGTGCTTTTCCCGCAGGTTGAAGC	164
Db	3161	TGGCGCATCTGGGCAATCCGGGAGAGAAAGACCCTGTGCTTTTCCCGCAGGTTGAAGC	3107
QY	165	GCATGAAAAGATTGTTCCGAGATGACTCTGTGTGATTTGACGTTGAGCGAAAACGACG	224
Db	3101	GCATGAAAAGATTGTTCCGAGAGCAGCTGTTCCTGATTTGACGTTGAGCGAAAACGACG	3042
QY	225	TTTACATGATGATTTGGGAGAGTGTGGCAGTCAAGCCTTTTAAACGATGAATGTTGCTT	284
Db	3041	TTACCATGATGATTTGGGAGAGTGTGGCAGTCAATGCTTTTAAACGATGAATGTTGCTT	2982
QY	285	CAGGCCACCTGGGATTCACAGTTCTGTGCGGCTTTTCCGAGACAGTTCGGATGGTATAC	344
Db	2981	CAGGCCACCTGGGATTCACAGCCCGTGTGAGACTGTTCCGAGACAGTTCGGATGGTATAGT	2922
QY	345	CCGAGCGCATCAGCAACCCGAGCAATACCGGCGACAGCCGGAATCTCCGTCCGAGTGTG	404
Db	2921	CCGAGCGCATCAGCAACCCGAGCAATACCGGCGACAGCCGGAATCTCCGTCCGAGTGTG	2862
QY	405	CAGATTAATGACAGCGGTGTGGGCGCTCGTGATTTACGTACGCGAGACGCGTATCTGTGC	464
Db	2861	CAGATTAATGACAGCGGTGTGGGCGCTCGTGATTTACGTACGCGAGACGCGTATCTGTGC	2802
QY	465	TGGATCCCGCGAAATGACATGATATCCCGTGAATTACCCGCGGGGCGCGCTGTTC	524
Db	2801	TGGATCCCGCGAAATGACATGATATCCCGTGAATTACCCGCGGGGCGCGCTGTTC	2742
QY	555	ATTCAACGTTTTTGAACCCGTGAGAGACGGGCGAGCTCGCGGTGTGCMAATGTGTTTACAGC	584
Db	2741	ATTCAACGTTTTTGAACCCGTGAGAGACGGGCGAGCCCGGGGTGCMAAATGTGTTTACAGC	2682
QY	585	GTGATGAGAGAGATGAAAGATGCTCCGACACGCTGTGCAGAACGCACT	631
Db	2681	GTGATGAGAGAGATGAAAGATGCTCCGACACGCTGTGCAGAACGCACT	

1993-07-22, Application US/10114170  
 Application No. US20030023075A1  
 SERIAL INFORMATION:  
 APPLICANT: Blattner, Frederick R.  
 Burland, Valerie  
 Perna, Nicole T.  
 Plunkett, Guy  
 Welch, Rod  
 TITLE OF INVENTION: No. US20030023075A1 Sequences of E. coli O157  
 NUMBER OF SEQUENCES: 265  
 CORRESPONDENCE ADDRESS:

1 APPLICATION NUMBER: 09/453,702  
2 FILING DATE: 03-DEC-1999  
3 APPLICATION NUMBER: 60/110,955  
4 FILING DATE: 04-DEC-1998  
5 ATTORNEY/AGENT INFORMATION:  
6 NAME: Seay, Nicholas J.



QY 525 ATTACGTTTTGAACCCGCTGGAGACCGGCGAGCTCGCGGNGCAATGTGTTTTACAGC 584  
|  
Db 16713 ATTACGCTTTTGAACCCGCTGGAGACCGGCGAGACCCGCGGNGCAATGTGTTTTACAGC 16654  
QY 585 GTGATGAGACAGATGAGATGCTCGACACGCTGCAAGAACGAGCT 631  
|  
Db 16653 GTGATGAGACAGATGAGATGCTCGACACGCTGCAAGAACGAGCT 16607

RESULT 12  
US-10-001-189-41  
; Sequence 41, Application US/10001189  
; Publication No. US20020173634A1  
; GENERAL INFORMATION:  
; APPLICANT: FRASER JR., MALCOLM J.  
; APPLICANT: LI, XU  
; APPLICANT: BEAM, TERESA  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TRANSPOSITION USING  
; TITLE OF INVENTION: MINIMAL SEGMENTS OF THE EUKARYOTIC TRANSFORMATION  
; FILE REFERENCE: 835910-92098  
; CURRENT APPLICATION NUMBER: US/10/001,189  
; PRIOR FILING DATE: 2001-10-30  
; PRIOR APPLICATION NUMBER: 60/244,984  
; PRIOR FILING DATE: 2000-11-01  
; PRIOR APPLICATION NUMBER: 60/244,677  
; NUMBER OF SEQ ID NOS: 70  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 41  
; LENGTH: 3662  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: pXL-Bac  
US-10-001-189-41

Query Match 73.4%; Score 519.2; DB 13; Length 3662;  
Best Local Similarity 98.5%; Pred. No. 6.9e-170;  
Matches 524; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 2 GATCCCATGCGTCAATTTTACGACATCTATTCTAGGGTTATCTAGTCATCAGCA 61  
|  
Db 983 GATCCCATGCGTCAATTTTACGACATCTATTCTAGGGTTATCTAGTCATCAGCA 61  
QY 62 TCATATCGTGGGCTTTTTCGGCTCACTGATCGCCCAAGCTGGGCTATCTGGGCA 1042  
|  
Db 1043 TCATATCGTGGGCTTTTTCGGCTCACTGATCGCCCAAGCTGGGCTATCTGGGCA 1102  
QY 122 CGGGAGAGAAAGCCCGTGCCTTTTCCCGGAGGTTGAAGCGGCAATGAAAGATTG 181  
|  
Db 1103 CGGGAGAGAAAGCCCGTGCCTTTTCCCGGAGGTTGAAGCGGCAATGAAAGATTG 1162  
QY 182 CGAGGATGACTGCTGTCATGACCTTGAGCGAAACGACGTTTACCATGATTCG 241  
|  
Db 1163 CGAGGATGACTGCTGTCATGACCTTGAGCGAAACGACGTTTACCATGATTCG 1222  
QY 242 GGAAGGTGGCCATGACGCTTTTAAAGGTGAAGTTCGTTGAGGCACTGGGATAC 301  
|  
Db 1223 GGAAGGTGGCCATGACGCTTTTAAAGGTGAAGTTCGTTGAGGCACTGGGATAC 1282  
QY 302 CAGTTGCTCGGCTTTTCCGGAACAAGTTCCGATGCTCAGCCCGAAGCGCATCAGCA 361  
|  
Db 1283 CAGTTGCTCGGCTTTTCCGGAACAAGTTCCGATGCTCAGCCCGAAGCGCATCAGCA 1342  
QY 362 CCGGAACATACCGGCGACAGCCGGAACGCGGTGCGGTTGAGATTAAGACGCG 421  
|  
Db 1343 CCGGAACATACCGGCGACAGCCGGAACGCGGTGCGGTTGAGATTAAGACGCG 1402  
QY 422 TGGGCGCTGGGATTTAGTCAAGGAGGACGGGTATCTGGCTGATGCGCGAGAATG 481  
|  
Db 1403 TGGGCGCTGGGATTTAGTCAAGGAGGACGGGTATCTGGCTGATGCGCGAGAATG 1462

QY 482 GACATGATACCCCGTGAGTTACCCGCGGCGCGCGCTGCTTATTCAGCTT 533  
|  
Db 1463 GACATGATACCCCGTGAGTTACCCGCGGCGCGCGCTGCTTATTCAGCTT 1514

RESULT 13  
US-10-001-189-53/c  
; Sequence 53, Application US/10001189  
; Publication No. US20020173634A1  
; GENERAL INFORMATION:  
; APPLICANT: FRASER JR., MALCOLM J.  
; APPLICANT: LI, XU  
; APPLICANT: BEAM, TERESA  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TRANSPOSITION USING  
; TITLE OF INVENTION: MINIMAL SEGMENTS OF THE EUKARYOTIC TRANSFORMATION  
; FILE REFERENCE: 835910-92098  
; CURRENT APPLICATION NUMBER: US/10/001,189  
; PRIOR FILING DATE: 2001-10-30  
; PRIOR APPLICATION NUMBER: 60/244,984  
; PRIOR FILING DATE: 2000-11-01  
; PRIOR APPLICATION NUMBER: 60/244,677  
; NUMBER OF SEQ ID NOS: 70  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 53  
; LENGTH: 4941  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: pXL-Bac-ECP  
US-10-001-189-53

Query Match 72.3%; Score 511; DB 13; Length 4941;  
Best Local Similarity 100.0%; Pred. No. 5.9e-167;  
Matches 511; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 CATGCGTCAATTTTACGACATCTATTCTAGGGTTATCTAGTCATCAGATCATA 66  
|  
Db 1129 CATGCGTCAATTTTACGACATCTATTCTAGGGTTATCTAGTCATCAGATCATA 1070  
QY 67 TCGACGGGCTTTTTCGGCTCACTGATCGCCCAAGCTGGGCTATCTGGGCA 126  
|  
Db 1069 TCGACGGGCTTTTTCGGCTCACTGATCGCCCAAGCTGGGCTATCTGGGCA 1010  
QY 127 AGGAAGAGCCCGTGCCTTTTCCCGGAGGTTGAAGCGGCAATGAAAGATTG 186  
|  
Db 1009 AGGAAGAGCCCGTGCCTTTTCCCGGAGGTTGAAGCGGCAATGAAAGATTG 950  
QY 187 ATGACTGCTGCTGATGACGTTTGAAGCGAAACGACGTTTACCATGATTCG 246  
|  
Db 949 ATGACTGCTGCTGATGACGTTTGAAGCGAAACGACGTTTACCATGATTCG 890  
QY 247 GTGTGCGCATGACGCTTTTAAAGGTGAAGTTCGTTGAGGCACTGGGATAC 306  
|  
Db 889 GTGTGCGCATGACGCTTTTAAAGGTGAAGTTCGTTGAGGCACTGGGATAC 830  
QY 307 CGTGGGCTTTTCCGGAACAAGTTCCGATGCTCAGCCCGAAGCGCATCAGCA 366  
|  
Db 829 CGTGGGCTTTTCCGGAACAAGTTCCGATGCTCAGCCCGAAGCGCATCAGCA 770  
QY 367 ACAATACCGGCGACAGCCGGAACGCGGTGCGGTTGAGATTAAGACGCGT 426  
|  
Db 769 ACAATACCGGCGACAGCCGGAACGCGGTGCGGTTGAGATTAAGACGCGT 710  
QY 427 CGCTGGGATTTAGTCAAGGAGGACGGGTATCTGGCTGATGCGCGAGAATG 486  
|  
Db 709 CGCTGGGATTTAGTCAAGGAGGACGGGTATCTGGCTGATGCGCGAGAATG 486  
QY 487 GGATACCCCGTGAGTTACCCGCGGCGCGCG 517  
|

Db 649 GGATACCCCGTAGATTACCGCGGCGCGC 619

## RESULT 14

US-10-001-189-51/c

; Sequence 51, Application US/10001189  
; Publication No. US20020173634A1

; GENERAL INFORMATION:

; APPLICANT: FRASER JR., MALCOLM J.

; APPLICANT: LI, XU

; APPLICANT: BEAM, TERESA

; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TRANSPOSITION USING

; TITLE OF INVENTION: MINIMAL SEGMENTS OF THE EUKARYOTIC TRANSFORMATION

; FILE REFERENCE: 835910-92098

; CURRENT APPLICATION NUMBER: US/10/001,189

; PRIOR FILING DATE: 2001-10-30

; PRIOR APPLICATION NUMBER: 60/244,984

; PRIOR FILING DATE: 2000-11-01

; PRIOR APPLICATION NUMBER: 60/244,677

; PRIOR FILING DATE: 2000-10-31

; NUMBER OF SEQ ID NOS: 70

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 51

; LENGTH: 4951

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: pXL-Bac-EYFP

US-10-001-189-51

Query Match 72.3%; Score 511; DB 13; Length 4951;

Best Local Similarity 100.0%; Pred. No. 5,9e-167;

Matches 511; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 7 CATCGTCAATTTTACGAGACTATCTTTAGGGTTAATCTAGCTGATCAGATCATA 66

Db 1129 CATCGTCAATTTTACGAGACTATCTTTAGGGTTAATCTAGCTGATCAGATCATA 1070

Qy 67 TCGTGGGCTCTTTTCCGGCTCAGTCAATCGCCCAAGCTGGCGCTATCTGGGCAATCGGG 126

Db 1069 TCGTGGGCTCTTTTCCGGCTCAGTCAATCGCCCAAGCTGGCGCTATCTGGGCAATCGGG 1010

Qy 127 AGAAGAAGCCCGGCTCTTTTCCGGCTCAGTCAATCGCCCAAGCTGGCGCTATCTGGGCAATCGGG 186

Db 1009 AGAAGAAGCCCGGCTCTTTTCCGGCTCAGTCAATCGCCCAAGCTGGCGCTATCTGGGCAATCGGG 950

Qy 187 ATGACTGCTGCTGATGACGTTGAGCGAAGCAAGCTTTAATCCATGATGATGGGAG 246

Db 949 ATGACTGCTGCTGATGACGTTGAGCGAAGCAAGCTTTAATCCATGATGATGGGAG 890

Qy 247 GTGAGCCATCAGCGCTTTTAAAGGTGAACTGTTGTTCAAGGCCACTGGGATACAGTT 306

Db 889 GTGAGCCATCAGCGCTTTTAAAGGTGAACTGTTGTTCAAGGCCACTGGGATACAGTT 830

Qy 307 CGTGGCGCTTTTCCGAGCAAGTTCCGGATGATCAGCCGGAAGCGATCAGCAACCGGA 366

Db 829 CGTGGCGCTTTTCCGAGCAAGTTCCGGATGATCAGCCGGAAGCGATCAGCAACCGGA 770

Qy 367 ACAATACCGGAGCAGCGGAACTGCGCGGATGACAGTTAATGACAGCGGTGCG 426

Db 769 ACAATACCGGAGCAGCGGAACTGCGCGGATGACAGTTAATGACAGCGGTGCG 710

Qy 427 CGCTGGATATTAAGTCAAGGAGACGGGATCTGAGTGCAGGAAATGACAT 486

Db 709 CGCTGGATATTAAGTCAAGGAGACGGGATCTGAGTGCAGGAAATGACAT 650

Qy 487 GGATACCCCGTAGATTACCGCGGCGCGC 517

Db 649 GGATACCCCGTAGATTACCGCGGCGCGC 619

## RESULT 15

US-10-001-189-52/c

; Sequence 52, Application US/10001189  
; Publication No. US20020173634A1

; GENERAL INFORMATION:

; APPLICANT: FRASER JR., MALCOLM J.

; APPLICANT: LI, XU

; APPLICANT: BEAM, TERESA

; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TRANSPOSITION USING

; TITLE OF INVENTION: MINIMAL SEGMENTS OF THE EUKARYOTIC TRANSFORMATION

; FILE REFERENCE: 835910-92098

; CURRENT APPLICATION NUMBER: US/10/001,189

; PRIOR FILING DATE: 2001-10-30

; PRIOR APPLICATION NUMBER: 60/244,984

; PRIOR FILING DATE: 2000-11-01

; PRIOR APPLICATION NUMBER: 60/244,677

; PRIOR FILING DATE: 2000-10-31

; NUMBER OF SEQ ID NOS: 70

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 52

; LENGTH: 4952

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: pXL-Bac-BGFP

US-10-001-189-52

Query Match 72.3%; Score 511; DB 13; Length 4952;

Best Local Similarity 100.0%; Pred. No. 5,9e-167;

Matches 511; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 7 CATCGTCAATTTTACGAGACTATCTTTAGGGTTAATCTAGCTGATCAGATCATA 66

Db 1129 CATCGTCAATTTTACGAGACTATCTTTAGGGTTAATCTAGCTGATCAGATCATA 1070

Qy 67 TCGTGGGCTCTTTTCCGGCTCAGTCAATCGCCCAAGCTGGCGCTATCTGGGCAATCGGG 126

Db 1069 TCGTGGGCTCTTTTCCGGCTCAGTCAATCGCCCAAGCTGGCGCTATCTGGGCAATCGGG 1010

Qy 127 AGAAGAAGCCCGGCTCTTTTCCGGCTCAGTCAATCGCCCAAGCTGGCGCTATCTGGGCAATCGGG 186

Db 1009 AGAAGAAGCCCGGCTCTTTTCCGGCTCAGTCAATCGCCCAAGCTGGCGCTATCTGGGCAATCGGG 950

Qy 187 ATGACTGCTGCTGATGACGTTGAGCGAAGCAAGCTTTAATCCATGATGATGGGAG 246

Db 949 ATGACTGCTGCTGATGACGTTGAGCGAAGCAAGCTTTAATCCATGATGATGGGAG 890

Qy 247 GTGAGCCATCAGCGCTTTTAAAGGTGAACTGTTGTTCAAGGCCACTGGGATACAGTT 306

Db 889 GTGAGCCATCAGCGCTTTTAAAGGTGAACTGTTGTTCAAGGCCACTGGGATACAGTT 830

Qy 307 CGTGGCGCTTTTCCGAGCAAGTTCCGGATGATCAGCCGGAAGCGATCAGCAACCGGA 366

Db 829 CGTGGCGCTTTTCCGAGCAAGTTCCGGATGATCAGCCGGAAGCGATCAGCAACCGGA 770

Qy 367 ACAATACCGGAGCAGCGGAACTGCGCGGATGACAGTTAATGACAGCGGTGCG 426

Db 769 ACAATACCGGAGCAGCGGAACTGCGCGGATGACAGTTAATGACAGCGGTGCG 710

Qy 427 CGCTGGATATTAAGTCAAGGAGACGGGATCTGAGTGCAGGAAATGACAT 486

Db 709 CGCTGGATATTAAGTCAAGGAGACGGGATCTGAGTGCAGGAAATGACAT 650

Qy 487 GGATACCCCGTAGATTACCGCGGCGCGC 517

Db 649 GGATACCCCGTAGATTACCGCGGCGCGC 619

Search completed: January 26, 2005, 17:51:21  
Job time : 381.34 secs

**This Page Blank (uspto)**